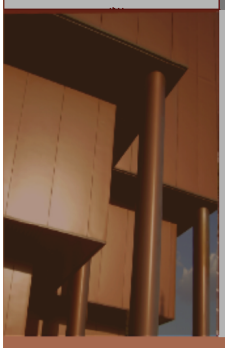
Tesis Doctoral

Ingeniería de Telecomunicación

Image analysis for diagnostic support in biomedicine: neuromuscular diseases and pigmented lesions.

Autor: Aurora Sáez Manzano
Directores: Begoña Acha Piñero
Carmen Serrano Gotarredona



Teoría de la Señal y Comunicaciones
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Tesis Doctoral
Ingeniería de Telecomunicación

Image analysis for diagnostic support in biomedicine: neuromuscular diseases and pigmented lesions.

Autor:

Aurora Sáez Manzano

Director:

Begoña Acha Piñero
Carmen Serrano Gotarredona
Profesor Titular

Teoría de la Señal y Comunicaciones
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Resumen

Esta tesis presenta dos sistemas implementados mediante técnicas de procesamiento de imagen, para ayuda al diagnóstico de enfermedades neuromusculares a partir de imágenes de microscopía de fluorescencia y análisis de lesiones pigmentadas a partir de imágenes dermoscópicas.

El diagnóstico de enfermedades neuromusculares se basa en la evaluación visual de las biopsias musculares por parte del patólogo especialista, lo que conlleva una carga subjetiva. El primer sistema propuesto en esta tesis analiza objetivamente las biopsias musculares y las clasifica en distrofias, atrofias neurógenas o control (sin enfermedad) a través de imágenes de microscopía de fluorescencia. Su implementación reúne los elementos propios de un sistema de ayuda al diagnóstico asistido por ordenador: segmentación, extracción de características, selección de características y clasificación. El procedimiento comienza con una segmentación precisa de las fibras musculares usando morfología matemática y una transformada Watershed. A continuación, se lleva a cabo un paso de extracción de características, en el cual reside la principal contribución del sistema, ya que no solo se extraen aquellas que los patólogos tienen en cuenta para diagnosticar sino características que se escapan de la visión humana. Estas nuevas características se extraen suponiendo que la estructura de la biopsia se comporta como un grafo, en el que los nodos se corresponden con las fibras musculares, y dos nodos están conectados si dos fibras son adyacentes. Para estudiar la efectividad que estos dos conjuntos presentan en la categorización de las biopsias, se realiza una selección de características y una clasificación empleando una red neuronal Fuzzy ARTMAP. El procedimiento concluye con una estimación de la severidad de las biopsias con patrón distrófico. Esta caracterización se realiza mediante un análisis de componentes principales. Para la validación del sistema se ha empleado una base de datos compuesta por 91 imágenes de biopsias musculares, de las cuales 71 se consideran imágenes de entrenamiento y 20 imágenes de prueba. Se consigue una elevada tasa de aciertos de clasificacion y se llega a la importante conclusión de que las nuevas características estructurales que no pueden ser detectadas por inspección visual mejoran la identificación de biopsias afectadas por atrofia neurógena.

La segunda parte de la tesis presenta un sistema de clasificación de lesiones pigmentadas. Primero se propone un algoritmo de segmentación de imágenes en color para ais-

lar la lesión de la piel circundante. Su desarrollo se centra en conseguir un algoritmo relacionado con las diferencias color percibidas por el ojo humano. Consiguiendo así, no solo un método de segmentación de lesiones pigmentadas sino un algoritmo de segmentación de propósito general. El método de segmentación propuesto se basa en un gradiente para imágenes en color integrado en una técnica de level set para detección de bordes. La elección del gradiente se derivada a partir de un análisis de tres gradientes de color implementados en el espacio de color uniforme CIE $L^*a^*b^*$ y basados en las ecuaciones de diferencia de color desarrolladas por la comisión internacional de iluminación (CIELAB, CIE94 y CIEDE2000). El principal objetivo de este análisis es estudiar cómo estas ecuaciones afectan en la estimación de los gradientes en términos de correlación con la percepción visual del color. Una técnica de level-set se aplica sobre estos gradientes consiguiendo así un detector de borde que permite evaluar el rendimiento de dichos gradientes. La validación se lleva a cabo sobre una base de datos compuesta por imágenes sintéticas diseñada para tal fin. Se realizaron tanto medidas cuantitativas como cualitativas. Finalmente, se concluye que el detector de bordes basado en la ecuación de diferencias de color CIE94 presenta la mayor correlación con la percepción visual del color.

A partir de entonces, la tesis intenta emular el método de análisis de patrones, la técnica de diagnóstico de lesiones pigmentadas de la piel más empleada por los dermatólogos. Este método trata de identificar patrones específicos, pudiendo ser tanto globales como locales. En esta tesis se presenta una amplia revisión de los métodos algorítmicos, publicados en la literatura, que detectan automáticamente dichos patrones a partir de imágenes dermoscópicas de lesiones pigmentadas. Tras esta revisón se advierte que numerosos trabajos se centran en la detección de patrones locales, pero solo unos pocos abordan la detección de patrones globales.

El siguiente paso de esta tesis, por tanto, es la propuesta de diferentes métodos de clasificación de patrones globales. El objetivo es identificar tres patrones: reticular, globular y empedrado (considerado un solo patrón) y homogéneo. Los métodos propuestos se basan en un análisis de textura mediante técnicas de modelado. En primer lugar una imagen demoscópica se modela mediante campos aleatorios de Markov, los parámetros estimados de este modelo se consideran características. A su vez, se supone que la distribución de estas características a lo largo de la lesión sigue diferentes modelos: un modelo gaussiano, un modelo de mezcla de gaussianas o un modelo de bolsa de características. La clasificación se lleva a cabo mediante una recuperación de imágenes basada en diferentes métricas de distancia. Para validar los métodos se emplea un conjunto significativo de imágenes dermatológicas, concluyendo que el modelo basado en mezcla de gaussianas proporciona la mejor tasa de clasificación. Además, se incluye una evaluación adicional en la que se clasifican melanomas con patrón multicomponente obteniendo resultados prometedores.

Finalmente, se presenta una discusión sobre los hallazgos y conclusiones más relevantes extraídas de esta tesis, así como las líneas futuras que se derivan de este trabajo.

Contents

CHAPTER 1

1 Introduction

Queda prohibido no sonreír a los problemas, no luchar por lo que quieres, abandonarlo todo por miedo, no convertir en realidad tus sueños.

Pablo Neruda

Biomedical imaging plays a central role in the global healthcare system as it contributes to early diagnosis in all major disease entities improving patient healthcare. However, diagnosis by imaging usually has a great subjective influence, since it is often based on the extraction of specific information and visual interpretation of medical image data. Therefore, factors such as difficulty in the perception of the image, visual feature extraction, viewing conditions, physician experience, fatigue or concentration at any time can lead to what is called diagnostic inconsistency.

In recent years there have been a number of systems designed to overcome these shortcomings. This type of system is commonly called CAD: Computer-aid diagnosis or computer aided diagnosis. Recent advances in these diagnostic support systems demonstrated that the application of digital image processing techniques increases the efficiency, diagnostic confidence and productivity of physicians, acting as a "second" opinion to the clinician.

CAD systems are fundamentally based on highly complex pattern recognition and their typical architecture includes: image pre-processing; segmentation for definition of region(s) of interest; features extraction and selection; and classification.

The main objective of this thesis is to research on different techniques within the image processing framework to solve several biomedical computing challenges, focusing mainly in two applications: neuromuscular diseases and pigmented lesions. Specifically, this thesis deals with the development of a computer-aided diagnosis system of neuromuscular diseases from fluorescence microscopic images of muscle biopsies and a computer-aided pattern classification system for pigmented lesions in dermoscopic images.

The main novelty in the computer-aided diagnosis system of neuromuscular diseases from fluorescence microscopic images is the application itself. There have not been found

works that address this problem in the literature. The most important contributions in the proposed system are:

- The use of a set of simple techniques to develop a complex system that achieves a successful solution to the problem raised.
- Extraction and study of structural features beyond human vision that help to improve the diagnosis of these diseases.

On the contrary, classification of pigmented lesion from dermoscopic images is a well-known problem. There are different melanoma diagnosis methods from dermoscopic images, among them pattern analysis is the method most commonly used for providing diagnostic accuracy. It seeks to identify specific patterns, which may be global and local. The novelty in the computer-aided pattern classification system proposed in this thesis is the development of novel algorithms, being the main contributions of this system:

- Development of a segmentation algorithm for colour images based on gradients correlated with the colour differences perceived by the human visual system.
- Development of global pattern classification methods based on modelling.

The thesis, therefore, is divided into two parts that address the two aforementioned applications. Chapter 2 thoroughly presents the computer-aided diagnosis system of neuromuscular diseases. This system includes the main steps constituting in a CAD system: segmentation of muscle fibres; feature extraction, both those that pathologists take into account to diagnose the diseases and structural features that the human eye cannot see, based on the assumption that the biopsy is considered as a graph, where the nodes are represented by each muscle fibre, and two nodes are connected if two fibres are adjacent; feature selection to study which features are more discriminant for the classification; and classification using a neural network together with a study of grading the severity of the diseases.

The second part, which is broken down into three chapters, deals with the computer-aided pattern classification system in pigmented lesions. Chapter 3 proposes a colour image segmentation algorithm. For its development we focus on the study of the perceived colour differences by the human visual system to attempt a segmentation algorithm as correlated with the human vision as possible, achieving, in this way, not only a successful pigmented lesion segmentation but a general purpose segmentation algorithm.

In Chapter 4 an extensive review of the works found in the literature that detect local or global patterns based on the pattern analysis method is presented.

Novel classification methods of global patterns based on modelling are proposed in Chapter 5. The modelling is applied in two senses: first a dermoscopic image is modelled as a Markov random field in the CIE$L^*a^*b^*$ colour space and the estimated parameters of this model are treated as features. Furthermore, the distribution of these features are supposed to follow different models along a lesion: a Gaussian model, a Gaussian mixture model and a bag-of-features histogram model.

Finally, in Chapter 6 conclusions and future lines are exposed.

PART I

Neuromuscular diseases

CHAPTER 2

Related Publication

- International journals:

  Sáez A., Rivas E., Montero-Sánchez A., Paradas C., Acha B., Pascual A., Serrano C., Escudero LM. Quantifiable diagnosis of muscular dystrophies and neurogenic atrophies through network analysis. *BMC Medicine.* 11(1), art. no. 77 (2013).

  Sáez A., Acha B., Montero-Sánchez A., Rivas E., Escudero LM., Serrano C. Neuromuscular disease classification system. *Journal of Biomedical Optics.* Jun;18(6):066017 (2013).

- International conferences:

  Sáez A., Serrano C., Acha B., Escudero LM. Segmentation of muscle fibres in fluorescence microscopy images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7325 LNCS (PART 2), pp. 465-472 (2012).

- Patent:

  Método para obtener información útil para el diagnóstico de enfermedades neuromusculares. Escudero L.M., Montero Sánchez A., Paradas C., Rivas E., Pascual A., Sáez A., Serrano C., Acha B. PCT/ES2012/070796

- Other communications:

  Sánchez-Gutiérrez D., Sáez A., Rivas E., Montero-Sánchez A., Paradas C., Acha B., Pascual A., Serrano C., Escudero L.M. Systems Biology methods for image analysis in development and disease. *Meeting Network Biology SIG of Meeting of the International Society for Computational Biology* (2013)

  Sáez A., Rivas E., Montero Sánchez A., Paradas C., Acha B., et. al.: Neuronuscular diseases diagnosis through computerized image analysis. *En Centro de Investigación Biomédica en Red, Enfermedades Neurodegenerativas CIBERNED* Diciembre 2011. Donostia (San Sebastian) (2011).

2 Neuromuscular disease classification system

Everything in the universe is interconnected... Thou canst not stir a flower, without troubling of a star.

FRANCIS THOMPSON

In this chapter the problem of classification of neuromuscular diseases is addressed. Diagnosis of neuromuscular diseases is based on subjective visual assessment of biopsies from patients by the pathologist specialist. A system for objective analysis and classification of muscular dystrophies and neurogenic atrophies through muscle biopsy images of fluorescence microscopy is presented. The procedure starts with an accurate segmentation of the muscle fibres using mathematical morphology and a watershed transform. A feature extraction step is carried out in two parts: 24 features that pathologists take into account to diagnose the diseases and 58 structural features that the human eye cannot see, based on the assumption that the biopsy is considered as a graph, where the nodes are represented by each fibre, and two nodes are connected if two fibres are adjacent are calculated. A feature selection step using sequential forward selection and sequential backward selection methods, a classification using a Fuzzy ARTMAP neural network, and a study of severity grading are performed on these two sets of features. A database consisting of 91 images was used: 71 images for the training step and 20 as testing set. No misclassification were obtained. It is concluded that the addition of features undetectable by the human visual inspection improves the categorization of atrophic patterns.

2.1 Neuromuscular diseases

Neuromuscular diseases (MND) are a group of more than 150 inherited or acquired neurological diseases affecting the muscles and nervous system. They are within the group of

so-called rare diseases. The diagnosis of neuromuscular diseases is mainly based on histological characterization and morphological evaluation of sections of muscle samples, usually skeletal muscle biopsies [Dubowitz et al., 2007].

Skeletal muscle constitutes voluntary muscle and it is innervated by the motor neurons of the somatic nervous system. Skeletal muscle consists of very long tubular cells, also called muscle fibres with diameters ranging between 10 and 100 microns and containing many nuclei located in the periphery of the cell. Muscular fibres are organized in fascicles, which are surrounded by a layer of connective tissue (named perimysium). Between the fibres within a fascicle appears the endomysium, a mesh of loose connective tissue composed of fine collagen and reticular fibres. Skeletal muscle structure is shown in Fig. 2.1. Muscle fibres can be of two types often called type I (slow twitch fibres) and type II (fast twitch fibres), both types of fibres arranged according to a disorganized patchwork along the fascicles [Helliwell, 1999].

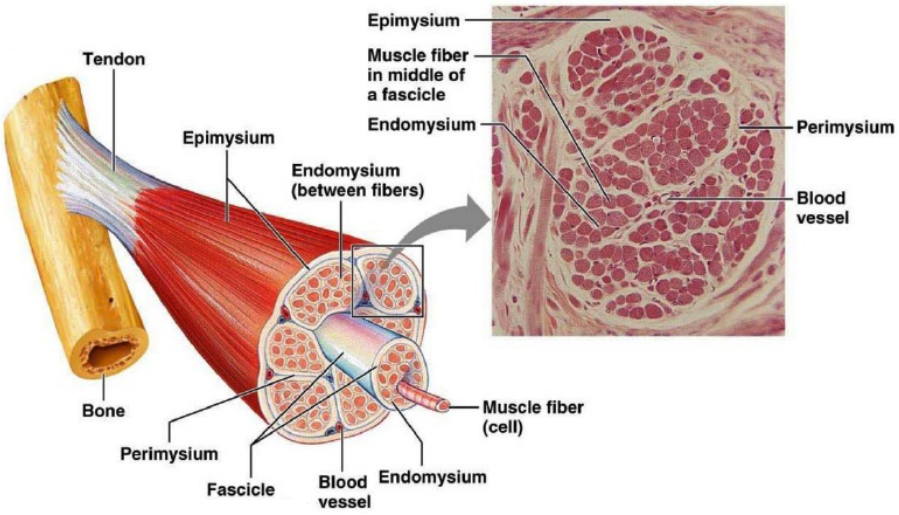


Figure 2.1 Structure of skeletal muscle.

Fibre size determination and its ratio between the two types of fibres is very important for diagnosis, as well as the size of collagen fibres, because the type of disease is reflected by a characteristic pattern of muscular section. Mainly there are three characteristic patterns:

- Dystrophic pattern: characterized by increased endomisio due to the appearance of fibrosis. Fibres acquire a more rounded morphology. Affectation is usually homogeneous within the muscle, affecting all muscles.

- No dystrophic Pattern: fibres also acquire more rounded morphology, although the endomisio does not increase. There is a greater disparity between the sizes of the fibres.

- Atrophic pattern: a large number of normal fibres at first glance, but small groups without specific pattern of small cells can appear. It may affect some muscles more than others. The affectation of fibres type I and type II can be different each other. Reinervations may occur, eliminating the mosaic. In advanced cases, early fibrosis may appear.

In this study, two neuromuscular diseases are analysed: muscular dystrophies (MD) (see Fig. 2.3 (e)) and neurogenic atrophies (NA) (see Fig. 2.3 (f)) and a combination of both patterns.

2.2 Antecedents

To study patient affections, the pathologist examines muscular biopsies of transversal sections under microscope focusing into a morphological muscular fibre analysis, a fundamental tool for the diagnosis of neuromuscular disorders [Dubowitz et al., 2007]. The evaluation of the changes in the morphological characteristics of a given biopsy with respect to the normal muscle is one of the main features for the diagnostic. However, the manual morphometric approach and interpretation of muscle biopsy material is a subjective, tedious and time-consuming task [Castleman et al., 1984].

Currently, tissue histopathology slide can be digitized and stored in digital image form. This progress has allowed to develop computer-aided diagnosis (CAD) for disease detection, diagnosis, and prognosis prediction in order to complement the opinion of the pathologist [Gurcan et al., 2009]. In this sense, in the recent literature works related to grading of prostate cancer [Doyle et al., 2007], detection of cervical cancer [Chan et al., 1996], classification of hepatocelular carcicoma [Huang and Lai, 2010], detection of cervical cell nuclei [Plissiti et al., 2011] or simple detection of different types of cells [Ficarra et al., 2011, Bergmeir et al., 2010, Harandi et al., 2010] can be found. Focusing into studies of muscular fibres we can find some works that address the segmentation of fibres in muscular biopsies [Todman and Claridge, 2000, Kim et al., 2007, Klemenčič et al., 1998], the classification of muscle fibre type [Sertel et al., 2011, Meunier et al., 2010, Karen et al., 2009, Behan et al., 2002] and the extraction of morphometric features [Garton et al., 2010]. However, studies of the characterization of neuromuscular disease based on image processing have not been found in the current literature.

2.3 Objective

The objective of this chapter is to develop a computer-aided diagnosis system of different neuromuscular diseases from fluorescence microscopy images of muscular biopsies.

To achieve this, a deep study of the features that better characterize the diseases is proposed. One of the novelties of the study is that it is based not only on the extraction of features related to the characteristics that the pathologist takes into account for diagnosis but on the search of new features with inherent properties that escape to the evaluation of the pathologist and could be more efficient for the classification of the different muscular images. In this sense, an extraction of morphometric and structural information based on

the assumption that the biopsy is considered as a graph, where the nodes are represented by each fibre and two nodes are connected if two fibres are adjacent, is proposed. To get the feature extraction an accurate segmentation is required. Mathematical morphology and a watershed transform are used to address this task. A fuzzy classification is carried out by a neural network architecture. And finally, it is proposed a study of the severity grading to analyse the results.

The proposed procedure is described in Fig. 2.2. As it is habitual in CAD systems, it consists of four main steps: segmentation, feature extraction, feature selection and classification.

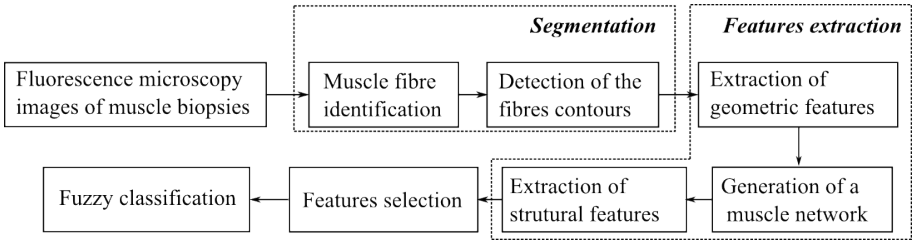


Figure 2.2 Flow diagram of the CAD system for neuromuscular diseases.

2.4 Muscle biopsy images

Muscular biopsies were processed by the standard methods of freezing and cutting with cryostat. The muscular fibre and the collagen content were detected by fluorescence microscopy. The antibodies Mouse anti-Myosin Heavy Chain (slow), Mouse anti-Myosin Heavy Chain (fast), Rabbit anti-Collagen Type VI were used using standard protocol for immunostaining. The sections were incubated with Alexa fluor 488 and Alexa fluor 568 secondary antibodies. All the slides were analysed under a fluorescence microscope (BX-61 Olympus with a DP70 camera) using a mercury lamp through a 470-490 nm or 560-579 nm band-pass filter to excite Alexa fluor 488 or Alexa fluor 568, respectively. The stained cells were photographed obtaining two high-resolution images (size of 4080 by 3072 pixels). A total of 91 images from 70 muscle biopsies stored in the Tissue Bank of the Hospital Universitario Virgen del Rocío, Seville, Spain, were processed.

An RGB image was created from the two images obtained, red component is the image obtained when the biopsy is excited at 560-579 nm. (Fig. 2.3 (b)), and green component is the image obtained when the biopsy is excited at 470-490 nm (Fig. 2.3 (c)). Fig. 2.3 (a) shows a sample of the resulting RGB image. Slow fibres in a dark colour, fast fibres in a reddish colour, collagen in a greenish colour and capillaries as small dark structures among the collagen can be observed.

All images have the same resolution (4080x3072 pixels). However, images shown in this Chapter represent only a part of the complete image (1150x1150 pixels) in order to visualize correctly the details.

Fig. 2.3 (d), (e) and (f) are examples of the variability between images. They belong

to patients with different ages and different diseases.

In this study 91 images from 70 subjects have been used; 41 control images, 27 dystrophy images and 23 atrophy images.

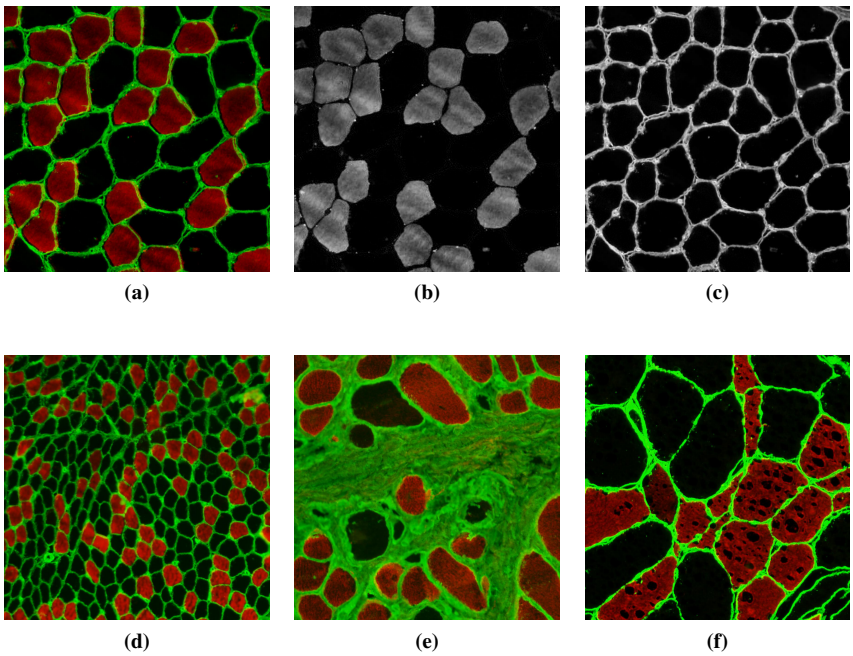


Figure 2.3 (a) Muscle biopsy image belonging to an adult. (b) R component. (c) G component. (d) Example of muscle biopsy image belonging to a child. (e) Example of muscle biopsy image affected by dystrophy. (f) Example of muscle biopsy image affected by atrophy.

2.5 Segmentation

A prerequisite to classify any disease is the ability to automatically identify structures present in the image [Gurcan et al., 2009]. Moreover, to achieve a robust morphological analysis, an accurate segmentation of the fibres in the biopsy image is required. Shape description and accurate segmentation is possible if the initial localization of the muscle fibres is known [Chen et al., 2012]. Therefore, the segmentation process is proposed to be divided into two steps: identification of the muscle fibre localization by applying morphological operators, and accurate detection of the fibre contours by watershed transformation.

It is important to note that the segmentation method must be automatic and valid to all types of images (see Fig. 2.3 (a), (d), (e) and (f)).

2.5.1 Fibre localization

The biopsy images we work with present different structures such us collagen, muscle fibres, capillaries and artefacts. The aim of this step is the correct identification of the muscle fibres.

The processing is performed on the G component of the image due to the high contrast between the muscle fibres and the collagen (see Fig. 2.3 (c)). Considering that the fibres are darker than surrounding collagen, intensity valleys in the image are searched. For this reason, the H-minima transform [Soille, 2003] is applied to the G component image in order to get homogeneous minima valleys. This transform has been successfully used in different medical applications [Jung and Kim, 2010], [Ali et al., 2011]. The H-minima or H-maxima transform is a powerful mathematical tool to suppress undesired minima or maxima. The h-minima transform reconstructs G from G+h, where h represents an arbitrary grey-level. By definition, the h-minima transform suppresses all regional minima whose depth is lower than or equal to the given h-value. Regional minima are connected components of pixels with a constant intensity value, and whose external boundary pixels have all a higher value.

The h-value has a direct influence on the number of segmented regions. The larger the h-value is, the fewer segmented regions are. As the darkest regions of the G component image represent the muscles fibres, intuitively, we can assume that the required h-value should be lower than the average intensity of the image. Three different values related to the average intensity were tested: two thirds of the average intensity of G (h_1), the half of the average intensity of G (h_2) and the third of the average intensity of G (h_3). They are calculated as:

$$h = K\frac{1}{N \cdot M}\sum_{i=1}^{N}\sum_{j=1}^{M} G(i,j), \qquad K = \frac{2}{3}, \frac{1}{2}, \frac{1}{3} \tag{2.1}$$

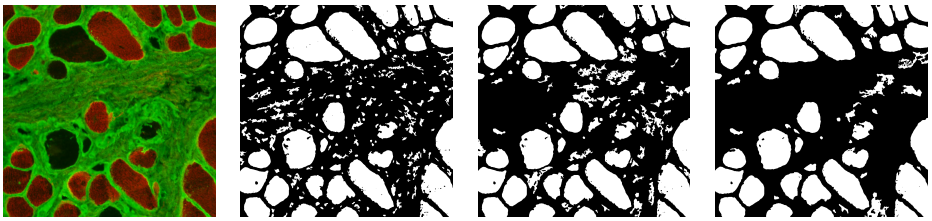
where $G(i,j)$ is the intensity value of G component at the pixel (i,j) and N,M are the image pixel dimensions.

The resulting image is a binary image, in which regions with pixel value of 1.0, displayed as white, represent candidate muscle fibres. Fig. 2.4 shows the influence of the three h-value on the number of segmented regions.

2.5.1.1 Results of the fibre localization

Results were qualitatively evaluated in only 10 1150x1150 pixel images. The reason of that is that the manual segmentation of each one is necessary in order to check the quality of the method and each image may have hundreds of cells. So, the manual delineation of them is a very time-consuming and tedious task. In Table 2.1 the number of cells segmented by the specialist is shown, as well as the number of the regions detected by the H-minima transform with the three different values of the h-value parameter. The number of detected regions should be equal or higher than the number of manually segmented cells due to the fact that two regions may join in one cell during the segmentation step. If the value is lower, it will involve loss of detected cells.

Although h1 provides a number of detected regions more similar to the number of cells



(a) (b) (c) (d)

Figure 2.4 (a) Muscle biopsy image. (b) H-minima transform with h-value=h1 (h1=60.85) . (c) H-minima transform with h-value=h2 (h2=45.63). (d) H-minima transform with h-value=h3 (h3=30.42).

Table 2.1 Number of manually segmented cells and number of detected regions by H-minima transform with different h-values. Bold values indicate that the number of detected regions is lower than the number of manually segmented cells, this will involve loss of detected cells..

Image	N manual cells	h1	h2	h3
1	77	**76**	79	83
2	103	108	128	144
3	40	250	295	366
4	45	256	530	1000
5	56	371	467	574
6	40	61	113	245
7	92	90	105	121
8	49	**48**	56	60
9	310	377	419	483
10	63	136	155	184

estimated by the pathologist (see Table 2.1), in some cases the number of detected regions is lower (see bold values in Table 2.1) meaning that some cells are not detected. Taking into account that the number shown here represents only a portion of the number of cells in the entire image, this number of lost cells could increase. Furthermore, the oversegmentation is due to the existence of artefacts and capillaries in the image. These regions can be removed by using morphological operators. For these reasons, the chosen h-value is h2 followed by a processing of the resulting image using mathematical morphology.

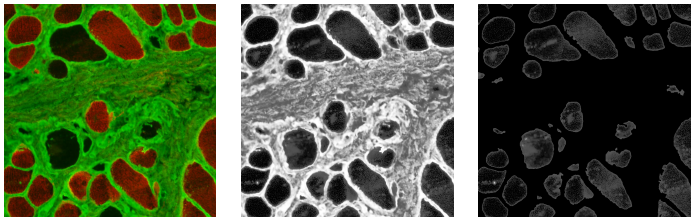
The morphological operators used in this step are summarized here. Regions with area smaller than 0.25% of the biggest dimension in pixels of the original image (taking into account the resolution indicated in Section 2.4) were removed by applying morphological opening [Soille, 2003]. Presumably, small regions correspond to capillaries. Irregularities (holes) within the regions detected were refilled by morphological reconstruction [Soille, 2003]. Finally, an erosion operator with a 3x3 pixel structural element was applied to prevent that adjacent cells were joined. The results of the number of regions

detected after this step are shown in the fourth column of Table 2.2.

Although the number of regions decreases, it would be desirable to adjust better the number of detected cells and the number of segmented cells. To this aim, two independent colour conditions were imposed to the detected regions in order to remove those that did not truly correspond to cells.

First condition: those regions whose average intensity of G component had a value higher than 25% of the maximum intensity value of G in the database were rejected as candidate fibre, as the muscle fibres present a low green value.

Second condition: a histogram equalization of the G component was performed (Fig. 2.5 (b)), and the average intensity values of the regions resulting after applying the first condition (Fig. 2.5 (c)) were computed. The regions whose average intensity value was higher than 90% of the maximum of all values were extracted. The resulting regions with R component value less than 20% of the R maximum intensity value in the database were removed.



(a) (b) (c)

Figure 2.5 (a) Muscle biopsy image. (b) Histogram equalization of the G component. (c) Intensity values of histogram equalization of the G component in the regions resulting after applying the first colour condition .

The thresholds mentioned were experimentally fixed and provided a correct segmentation of the 91 images analysed.

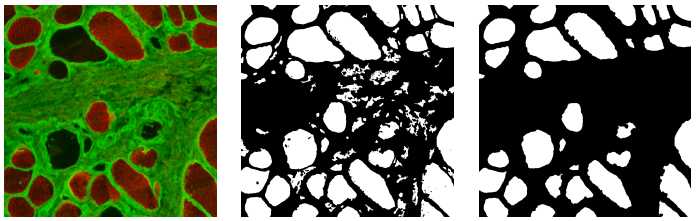
With the application of the mentioned morphological operators and the two colour conditions capillaries and artefacts present among the collagen are removed (Fig. 2.6 (c)).

The fifth column of Table 2.2 presents the final results.

2.5.2 Detection of fibre contours

Once the localizations of the muscle fibres are identified, an accurate detection of their contours is required for a later robust morphological analysis. For this objective, two well-known techniques were applied: level set and watershed transform. Both methods are briefly explained below.

Level set methods [Osher and Sethian, 1988] have been widely used as a global approach towards the optimization of active contours for the segmentation of objects of



(a) (b) (c)

Figure 2.6 (a) Muscle biopsy image. (b) Mask image resulting after applying the H-minima transform with h-value=h2 (h2=45.63). (c) Mask image resulting after applying the morphological operators and the two colour conditions .

Table 2.2 Number of manually segmented cells and number of detected regions by H-minima transform with h-value=half of the average intensity of G after application of morphological operators and two colour conditions.

Image	N cells manual	h2	Morphological operators	colour conditions
1	77	79	77	77
2	103	128	103	103
3	40	295	59	40
4	45	530	104	48
5	56	467	59	56
6	40	113	47	41
7	92	105	91	91
8	49	56	52	49
9	310	419	333	310
10	63	155	74	65

interest from the background. In the literature numerous works related to this technique have been proposed. In this thesis, the method developed by Li et al. is used (explained in detail in Section 3.3) [Li et al., 2005]. The initial curve required is the contour of the binary image resulting of the muscle fibre localization detection, explained in the previous section.

In the watershed procedure [Meyer and Beucher, 1990], an image is viewed as a topographic surface: the higher the value of a pixel, the higher the altitude at the corresponding point on the topographic surface or relief. The watershed transform is usually applied to the gradient image. The minima in the gradient image will correspond to sites within homogeneous regions in the original image. However, the watershed algorithm yields results with substantial over-segmentation; that is, the number of segmented regions could be much larger than desired, with regions being broken into multiple smaller regions. This undesirable result is due to the fact that the gradient image used in the process is sensitive to noise. The problem of over-segmentation can be overcome with the use of markers that identify the objects. The object contours in the gradient image can be seen

as the highest crest-lines around the object markers. In our case, the image gradient is calculated in the G component and the internal and external markers are derived from the fibre localization step (Section 2.5.1). The binary image resulting of Section 2.5.1 constitute the internal makers. Whereas the external markers were extracted by applying another watershed transform to the binary image used as internal markers.

Fig. 2.7 shows the steps followed in the fibre segmentation.

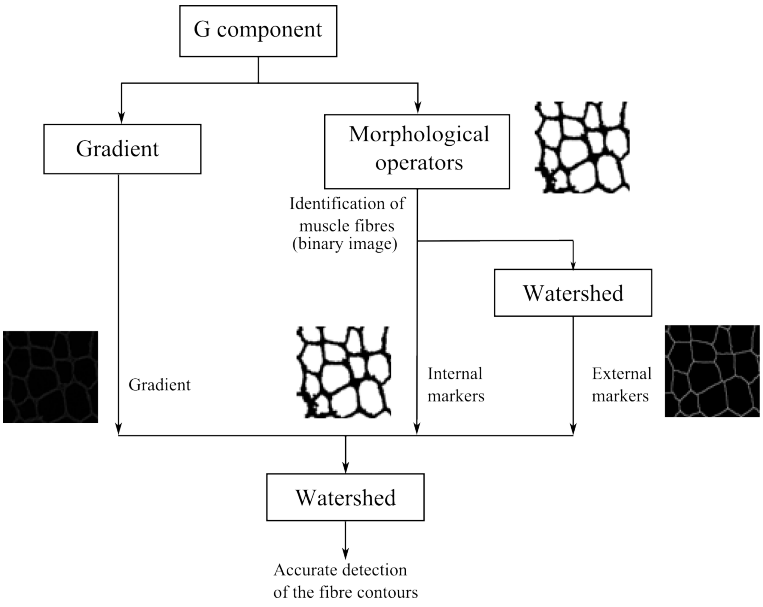


Figure 2.7 Steps followed in the fibre segmentation.

2.5.2.1 Results of the detection of the fibre contours

To evaluate the performance of both methods, 10 images manually segmented by the specialist were used. The results were evaluated with the Jaccard coefficient and the Dice coefficient. The Jaccard index, also known as the Jaccard similarity coefficient, is a statistical parameter used to compare the similarity and diversity of sample sets. It is defined as the size of the intersection divided by the size of the union of the segmented cells. Assuming two sets corresponding to the set of pixels manually segmented (Pm) and pixels automatically segmented (Ps), Jaccard's coefficient is defined as:

$$J = \frac{|P_m \cap P_s|}{|P_m \cup P_s|} \tag{2.2}$$

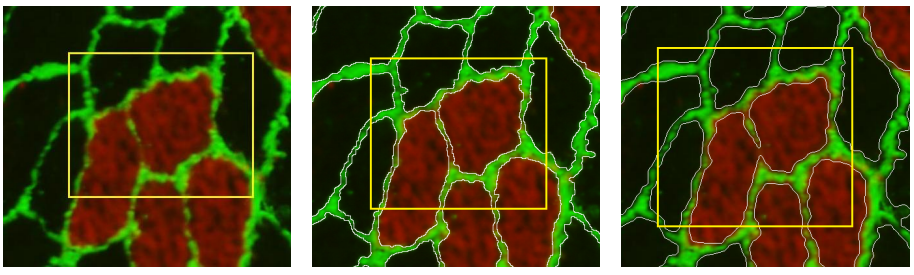
The Dice coefficient, D, is also a similarity measure defined as:

$$D = \frac{2|P_m \cap P_s|}{|P_m| + |P_s|} \tag{2.3}$$

Results are presented in Table 2.3. As it can be seen, the watershed transform outperforms level set technique in all the images. Examples of both segmentation are shown in Fig. 2.8. It is important to note that not only the accuracy in detecting the contours of the fibres is higher in watershed (Fig. 2.8 (b)) than in level sets (Fig. 2.8 (c)), but also the watershed transform is able to separate two independent cells although they seem to be linked (see yellow rectangles).

Table 2.3 Segmentation results for the watershed transform and level set segmentation, evaluated by the Dice coefficient and the Jaccard coefficient. Bold values indicate best results. .

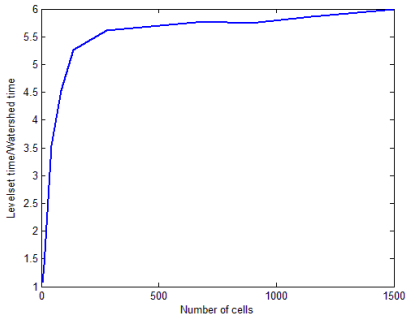
Image	Dice coefficient		Jaccard index	
	Watershed	level sets	Watershed	level sets
1	0.963	0.94	0.927	0.8969
2	0.966	0.93	0.933	0.868
3	0.97	0.95	0.942	0.912
4	0.969	0.952	0.939	0.908
5	0.969	0.949	0.94	0.903
6	0.956	0.938	0.915	0.881
7	0.96	0.917	0.922	0.843
8	0.975	0.947	0.951	0.89
9	0.971	0.917	0.94	0.84
10	0.973	0.922	0.948	0.854
Average	**0.967**	0.936	**0.934**	0.879



(a) (b) (c)

Figure 2.8 (a) Muscle biopsy image. The yellow rectangle indicates linked cells. (b) Watershed segmentation result. The yellow rectangle indicates a good result. (c) Level set segmentation result. The yellow rectangle indicates a bad result because two linked cells have been segmented as only one.

Furthermore, the computational cost is also lower for the watershed algorithm. Fig. 2.9 shows that the ratio of computational time between both segmentations (time consumed by the level sets algorithm /time consumed by watershed algorithm) increases logarithmically with the number of cells to be segmented in the image. In addition, it should be noted that the time spent on segmenting the images by the proposed method is significantly less than that needed by the specialist (see Table 2.4).



(a)

Figure 2.9 Ratio between computational cost of both segmentation methods and number of cells segmented.

Table 2.4 Time spent on segmenting the images by the proposed method and by the specialist.

Image	Manual segmentation	Proposed method
1	20 min.	33 sec.
2	33 min.	35.5 sec.
3	14 min.	33 sec.
4	22 min.	33.7 sec.
5	11 min.	30.14 sec.
6	8 min.	30 sec.
7	29 min.	41 sec.
8	15 min.	34 sec
9	40 min.	45 sec.
10	20 min.	33 sec.
Average	21.2 min	31.5 sec.

2.6 Feature extraction

To identify if a biopsy is affected by a pathology, an objective analysis of the biopsy is needed. The morphological and structural characteristics of the whole biopsy constitute a vector, also called biosignature [Chen et al., 2012]. The objective of this section is to extract features that correspond to the visual attributes defined by clinicians as particularly important for the diagnosis of the mentioned pathology and its grading of severity, as well as a study of new features with inherent properties that escape to the evaluation of the pathologist and that could be useful for the characterization of these diseases. In this sense, both morphological features and structural features are proposed.

2.6.1 Morphological features

The morphological characteristics can be described by shape or geometry [Chen et al., 2012]. Formulation of morphological features is an easy and fast way to automatize the manual morphological quantification, which is very laborious and subjective.

Features such as the area of each fibre, the average of the areas of type I (slow) and type II (fast) fibres or the major and the minor axis lengths of the cells are some of the features calculated. Type II fibres are identify by a high intensity average of R-component. In Table 2.5 the 14 characteristics computed are shown.

Table 2.5 14 morphological features of the cells.

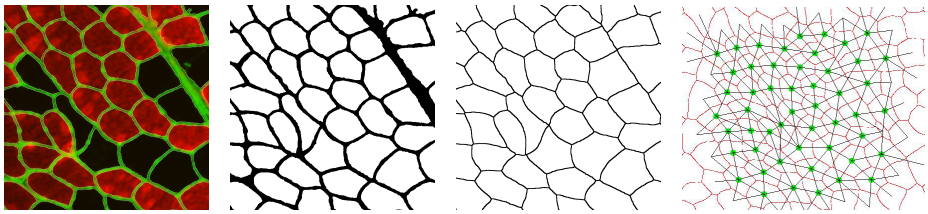
1	Average Area	8	Average minor Axis
2	Std. Dev. Area	9	Average ratio Axis
3	Average Area of slow cells	10	Std. Dev. ratio Axis
4	Std. Dev. Area of slow cells	11	Average Convex Hull
5	Average Area of fast cells	12	Std. Dev. Convex Hull
6	Std. Dev. Area of fast cells	13	Average Angles
7	Average Major axis	14	Std. Dev. Angles

2.6.2 Structural features

To address this approach each biopsy image was interpreted as a graph. Graphs are efficient data structures to represent spatial data and an effective way to represent structural information by defining a large set of topological features [Gurcan et al., 2009]. Formally, a simple graph $G = (V, E)$ is an undirected and unweighed graph without self-loops, with V and E being the node and edge set of graph G, respectively. Applying this concept to our problem we generate a cellular network, in which the muscle fibres are represented by nodes and two nodes are connected if two fibres are adjacent.

To identify the neighbourhood of each fibre we generate a mosaic such that each fibre contour is expanded to reach the expanded contour of the adjacent fibre. This concept was addressed by applying a watershed transform to the binary image resulting from fibre

detection step (Section 2.5.2). An example is shown in Fig. 2.10 (c).



(a) (b) (c) (d)

Figure 2.10 (a) Muscle biopsy image.(b) Mask of the watershed segmentation result (c) Mosaic, where each fibre contour is expanded to reach the expanded contour of the adjacent fibre. d) Graph, where each node is represented by the mass centre of each fibre and the neighbourhood relations are mapped into the edges .

It should be noted that an important feature was extracted from this mosaic. The ratio between the area of a fibre (A2) and the area when its contour is expanded (A1) (features 15 and 16 in Table 2.6), is an indicative of the amount of the collagen, and therefore an indicative of the existence of fibrosis, one of the main attributes of the muscular dystrophies.

In this point, a vector of 24 features (included the 14 geometrical ones) was fixed. Characteristics such as the number of neighbours or in particular, the number of neighbours of each type of fibre were added. Table 2.6 shows the features added. These 24 features try to emulate the characteristics that the pathologist take into account for diagnosis

Table 2.6 Structural features.

15	Average Ratio A1/A2	20	Std. Dev. Neighbours of fast fibres
16	Std. Dev. Ratio A1/A2	21	Slow Neighbours of slow fibres
17	Average Neighbours	22	Fast Neighbours of slow fibres
18	Std. Dev. Neighbours	23	Slow Neighbours of fast fibres
19	Std. Dev. Neighbours of slow fibres	24	fast Neighbours of Fast fibres

However, another set of features that escape to the evaluation of human vision is extracted. For this purpose a weighted graph derived from the muscular biopsy was generated, where each node was represented by the mass centre of the each fibre and the neighbourhood relations are mapped into weighted edges, where each weight corresponds to the Euclidean distance between nodes (see Fig. 2.10 (d)).

58 new characteristics were incorporated. The first 14 characteristics out of the 58 ones were obtained from the parameters already computed (area, axis, convex hull, angles

and ratio A1/A2) but taking into account the neighbourhood of each fibre, i.e. the ratio between the parameters and the average values within a neighbourhood constituted these new 14 features (features 25-38 in Table 2.7). The 44 remaining features are computed from graph theory (features 39-82 in Table 2.7), when it is applied to an undirected and weighted graph. Table 2.7 shows these 58 new features. Some definitions are given below:

Strength of a node: sum of weights of links connected to the node. The weights are defined as the distance in pixels between adjacent nodes.

Clustering coefficient of a node: probability that the neighbours of this vertex (all other vertices to which it is connected by an edge) are also connected to each other.

Eccentricity of a node: the maximal shortest path length between a node and any other node.

Betweenness centrality of a node: the number of shortest paths from all vertices to all others that pass through that node

Radius of a graph: the minimum eccentricity

Diameter of a graph: the longest of all shortest paths that is the maximum eccentricity, that is the maximum eccentricity.

Efficiency of a graph: the inverse of the average shortest path length in the network

Pearson: Pearson correlation calculation applied to each pair of linked nodes. The result always lies in the range [-1 1] with a negative result indicating that nodes of dissimilar degree tend to be linked and a positive result indicating that nodes of similar degree tend to be linked.

Algebraic connectivity: the second smallest eigenvalue of the Laplacian matrix. The Laplacian matrix contains the node degree as diagonal elements, and -1 for all cells corresponding to existing edges and 0 for cells corresponding to absent edges.

S metric: the sum of products of degrees across all edges.

Assortativity: correlation coefficient between the degrees of all nodes on two opposite ends of a link. A positive assortativity coefficient indicates that nodes tend to link to other nodes with the same or similar degree.

Density: the fraction of present connections to possible connections. Connection weights are ignored in calculations.

Transitivity: it measures the probability that the adjacent nodes of a node are connected.

Modularity: it is a statistic that quantifies the degree to which the network may be subdivided into such clearly delineated groups.

To avoid errors due to the lack of neighbours of the fibres at the image border, the characteristics are calculated on a region of interest (ROI) chosen by the users, such that at least one row of fibres around the ROI exists (see 2.10 d)).

Table 2.7 58 new structural features.

25	Average Ratio Neighbours Area	54	Std. Dev. Eccentricity of fast cells
26	Std. Dev. Ratio Neighbours Area	55	Average Eccentricity of slow cells
27	Average Ratio Neighbours major axis	56	Std. Dev. Eccentricity of slow cells
28	Std. Dev. Ratio Neighbours major axis	57	Average Betweenness centrality
29	Average Ratio Neighbours minor axis	58	Std. Dev. Betweenness centrality
30	Std. Dev. Ratio Neighbours minor axis	59	Average Betweenness centrality of fast cells
31	Average Ratio Neighbours ratio axis	60	Std. Dev. Betweenness centrality of fast cells
32	Std. Dev. Ratio Neighbours ratio axis	61	Average Betweenness centrality of slow cells
33	Average Ratio Neighbours convex hull	62	Std. Dev. Betweenness centrality of slow cells
34	Std. Dev. Ratio Neighbours convex hull	63	Average Shortest paths lengths
35	Average Ratio Neighbours angles	64	Std. Dev. Shortest paths lengths
36	Std. Dev. Ratio Neighbours angles	65	Average Shortest paths lengths from fast cells to fast cells
37	Average Ratio Neighbours ratio A1/A2	66	Std. Dev. Shortest paths lengths from fast cells to fast cells
38	Std. Dev. Ratio Neighbours ratio A1/A2	67	Average Shortest paths lengths from fast cells to slow cells
39	Average Strengths	68	Std. Dev. Shortest paths lengths from fast cells to slowcells
40	Std. Dev. Strengths	69	Average Shortest paths lengths from slow cells to slow cells
41	Average Strengths of fast cells	70	Std. Dev. Shortest paths lengths from slow cells to slow cells
42	Std. Dev. Strengths of fast cells	71	Average Shortest paths lengths from slow cells to fast cells
43	Average Strengths of slow cells	72	Std. Dev. Shortest paths lengths from slow cells to fast cells
44	Std. Dev. Strengths of slow cells	73	Radius
45	Average Clustering coefficient	74	Diameter
46	Std. Dev. Clustering coefficient	75	Efficiency
47	Average Clustering coefficient of fast cells	76	Pearson correlation
48	Std. Dev. Clustering coefficient of fast cells	77	Algebraic connectivity
49	Average Clustering coefficient of slow cells	78	S metric
50	Std. Dev. Clustering coefficient of slow cells	79	Assortativity
51	Average Eccentricity	80	Density
52	Std. Dev. Eccentricity	81	Transitivity
53	Average Eccentricity of fast cells	82	Modularity

2.7 Feature selection

Feature selection has two benefits. It reduces the cost of data collection and computational cost of recognition, and it usually improves the generalization performance of the classifier. Actually, a large set of features may possibly be detrimental to the classification performance, a phenomenon known as the curse of dimensionality. Feature selection is a means to select the relevant and important features from a large set of features. An optimal feature selection method would require an exhaustive search, which is not practical for a large set of features generated from a large dataset. Therefore, several heuristic algorithms have been developed, which use classification accuracy as the optimality criterion [Gurcan et al., 2009].

The well-known feature selection methods namely sequential forward selection (SFS) and sequential backward selection (SBS) [Keinosuke, 1990] are used in this work. SFS works by sequentially adding the feature that most improves the classification performance; similarly, SBS begins with the entire feature set and sequentially removes the feature that less improves the classification performance. While these methods still cannot guarantee optimality of the selected feature subset, they have been shown to perform very well compared to other feature selection methods [Jain and Zongker, 1997] and are, furthermore, much more computationally efficient [Pudil et al., 1994].

As it has been mentioned, these methods use classification accuracy as the optimality criterion. In this case the classification was performed by a Fuzzy-ARTMAP Neural network. It is a neural network architecture developed by Grossberg and Carpernter [Carpenter et al., 1992] and it is based on Adaptive Resonance Theory (ART). Fuzzy-ARTMAP is a supervised learning classification architecture for analogue-value input pairs of patterns where each individual input is mapped to a class label.

To evaluate the classification accuracy, a set of 71 images was used. 34 control biopsies from quadriceps and biceps, 20 biopsies images affected by dystrophy and 17 biopsies images affected by neurogenic atrophy. Three studies were carried out;

- Comparison between the three groups of images (controls (C) - muscular dystrophy (MD)- neurogenic atrophy (NA)),
- Comparison between control and dystrophies (biopsies of muscle affected by dystrophy belong to quadriceps)
- Comparison between control and neurogenic atrophies (biopsies of muscle affected by atrophy belong to biceps)

For each comparison the selection performance was evaluated by four-fold cross validation (XVAL) [Kohavi, 1995]. In this sense, the disadvantage of sensitiveness to the order of presentation of the training set, that the SBS and SFS methods present, was diminished. To perform the XVAL method four disjoint subsets of each class (control, dystrophy, neurogenic atrophy) was used. Three of these subsets served as training set for the neural network, while the other one was used as validation set. Then, the procedure was repeated interchanging the validation subset with one of the training subsets, and so on till the four subsets were used as validation sets. The final classification error was calculated as the average of the errors for each XVAL run.

Table 2.8 Feature selection results for the first feature set.

Comparison	24 features	
	Selected features	Class. error
Controls-dystrophies (C/MD)	19 18 15	0%
Controls-atrophies (C/NA)	12 20 21 22	0%
Controls-dystrophies-atrophies (C/MD/NA)	20 9 19 18 16 21 14 17	13.22%

Table 2.9 Feature selection results for the second feature set.

Comparison	82 features	
	Selected features	Class. error
Controls-dystrophies (C/MD)	25 9	0%
Controls-atrophies (C/NA)	34 16 21 45	0%
Controls-dystrophies-atrophies (C/MD/NA)	25 32 14 16 58 62 15 30 26	1.48%

2.7.1 Results of feature selection

The feature selection procedure was performed twice for the three comparisons mentioned. The first selection was carried out on the 24 first features described in the previous Section (see Table 2.5 and 2.6), and the second selection was performed on the 82 features (see Table 2.7). The results are shown in Table 2.8 and Table 2.9 for both feature sets, in which the selected features and the classification error obtained with this selection are shown. The selected features are presented in descending order of discrimination power.

As it can be seen, when we compare between both quadriceps-dystrophies and biceps-atrophies, the classification success is 100%, therefore, adding new features does not improve the classification error in this stage of training. In the following Section we will study what classification error is obtained when we classify new biopsies, that were not included in the stage of feature selection. However, in the case of the distinction between the three categories, the classification error decreases when we add structural features. In the next Section we check how these sets of features are good to classify new biopsies.

2.8 Classification

The extracted features represent the inputs to a classification procedure. The classifier used in this application is a Fuzzy ARTMAP neural network [Carpenter et al., 1992]. It consists in a supervised learning classification architecture for analog-value input, each

individual input is mapped to a class label.

The fuzzy ARTMAP system incorporates two fuzzy ART modules, ART_a and ART_b, that are linked together via an inter-ART module, F^{ab}, called a map field (see Fig. 2.11).

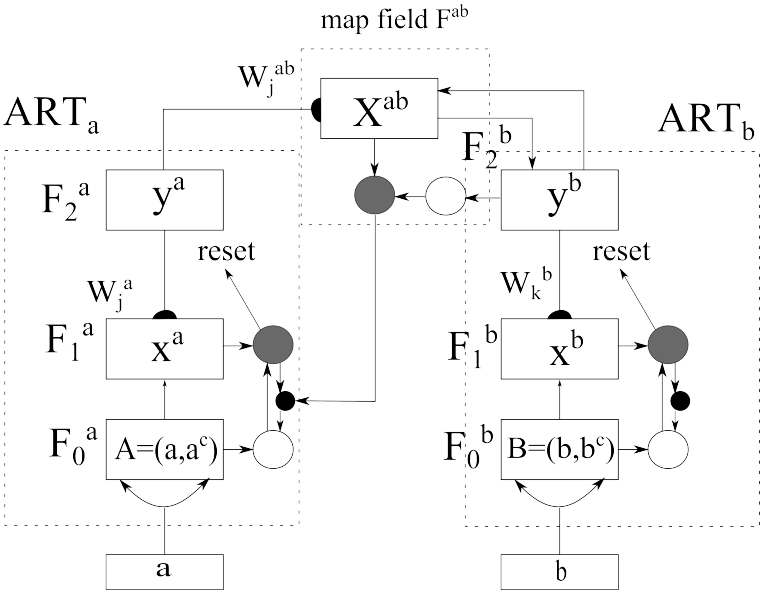


Figure 2.11 Fuzzy ARTMAP architecture.

Each ART system includes a field, F_0, of nodes that represent a current input vector and a field, F_1, that receives input from F_0. F_2^a is composed by N j nodes (j=1,..,N, N=number of training images), W_j^a denotes the weight vector that associates the j node of F_2^a to F^{ab} map. F_2^b is composed by M k nodes (k=1,..,M, M= number of the classes), W_k^b denotes the weight vector that associates the k category of F_2^b to F^{ab} map. W_{jk}^{ab} are the weights that emanate from every node j to every node k.

In the prediction stage, “a" (features) is the input vector, and it is hoped that the system responds with the “b" vector (classes). In this stage, there will be activated a J node of F_2^a. This J node is activated throughout the choice function T_j:

$$T_j(A) = \frac{|A \wedge W_j^a|}{\alpha + |W_j^a|} \tag{2.4}$$

where the fuzzy AND operator ($\wedge$) is defined by:

$$(p \wedge q)_i \equiv min(p_i, q_i) \tag{2.5}$$

The activated J node fulfils:

$$T_J = max\{T_j : j = 1...N\} \tag{2.6}$$

The W_{jk}^{ab} associated to this node, activates a K node of F_2^b. This K node is the class chosen that the system predicts associated to the 'a' input. However, in our implementation, the category choice is modified. As each node j of F_2^a is associated a training image, we have taken into account the average of the values T_j by categories. Let cl (cl=1,..,M) the category (M= number of the categories), in our case, cl can be control, dystrophy or atrophy. Let n_cl the number of the training images of the cl category. The category choice procedure is:

for cl=1:M
k_{cl}=mean(T_j), j=1:n_{cl}
end

The category chosen fulfils:

$$K_{CL} = max\{k_{cl} : cl = 1 : M\} \tag{2.7}$$

The M k_{cl}'s computed for each input, allows to present a classification called fuzzy, since for each input we obtain a value for each possible category. The system assigns the category (K_{CL}) corresponding to the highest value k_{cl}, however, it also provides an idea of how of reliable is the output. It can happen that for a biopsy (input) the response of the system does not provide a clear high value in the k_{cl}'s and by the contrary provides two similar values for two different categories. This fact implies that the biopsy can present a combination of different pathologies or a preliminary stage of a pathology. Some examples of this concept are presented in the classification results.

2.8.1 Classification results

To evaluate the classification procedure, 20 test images were used; 7 controls, 7 dystrophy, 5 neurogenic atrophy and 1 atrophy with a pseudo-dystrophic component. These images were not part of the training set to select the most discriminant features. Table 2.10 and Table 2.11 show the classification results for the features selected in the previous Section (Table 2.8 and Table 2.9).

Table 2.10 Results of classification 20 images; 7 controls, 7 dystrophies, 5 neurogenic atrophies and 1 atrophy of pseudo-dystrophic nature when the first feature set is used.

Comparison	24 features Class. error
Controls-dystrophies (C/MD)	100% (15 images/15 images)
Controls-atrophies (C/NA)	92.31% (12 images/13 images)
Controls-dystrophies-atrophies (C/MD/NA)	85% (17 images/20 images)

As regards to Table 2.10 and Table 2.11, both in the classification between C/NA and C/MD/NA, all errors were made in atrophies that were classified as controls. This fact is due to the low degree of severity that the biopsies affected by atrophy present and,

Table 2.11 Results of classification 20 images; 7 controls, 7 dystrophies, 5 neurogenic atrophies and 1 atrophy of pseudo-dystrophic nature when the second feature set is used.

Comparison	82 features Classification error
Controls-dystrophies (C/MD)	100% (15 images/15 images)
Controls-atrophies (C/NA)	100% (13 images/13 images)
Controls-dystrophies-atrophies (C/MD/NA)	85% (17 images/20 images)

therefore, to the great similarity with control biopsies. For the case C/MD/NA, where the three categories are compared, the classification error made with both sets of features is the same (85%) although in the feature selection procedure the classification error was considerably less for the case of 82 features (see Tables 2.8 and 2.9).

Otherwise, the numeric values (k_{cl}) obtained as output of the neural network could be considered as indicative of the degree of certainty of belonging to a class. In Table 2.12 an example of the k_{cl} values for a control biopsy image, a dystrophy with 2 degree of severity, an atrophy with 0-1 degree of severity and an atrophy with pseudo-dystrophic component are shown. The cases whose k_{cl}s differ less than 0.15 will require a further study of the biopsy.

For these cases we can give a more reliable result if we apply the classification criterion presented in Table 2.13. The procedure is to classify each biopsy with the three training cases of the neural network: C/MD/NA, C/MD and C/NA. Depending on the category selected by each neural network, the biopsy is classified as Table 2.13 indicates.

Table 2.12 Output values of the Fuzzy ARTMAP trained with controls-dystropies-atrophies with the 82 features, when it classifies different biopsies. Bold values in the two first cases indicate the highest kcl values obtained. This means that the biopsy belongs to this class. For the third and fourth cases there is no value predominantly higher than others. These cases present two similar values for different classes (kcl) indicated in bold. This means that a further study of the biopsy will be required..

Biopsies	Fuzzy ARTMAP output		
	k_1 (Controls)	k_2 (Dystrophy)	k_3 (Atrophy)
Control	**0.904**	0.757	0.810
Dystrophy with 2 degree of severity	0.631	**0.813**	0.750
Atrophy with 0-1 degree of severity	**0.867**	0.795	**0.855**
Atrophy with pseudo-dystrophic nature	0.688	**0.860**	**0.850**

Table 2.13 Classification criterion.

C/MD/NA output	C/MD output	C/NA output	Final categorization
C	C	C	C
C	C	NA	C or low atrophic (requires a study)
NA	C	NA	NA
MD	MD	C	MD
MD	MD	NA	Possible atrophy with dystrophic nature (requires a study)
NA	MD	NA	Possible atrophy with dystrophic nature (requires a study)

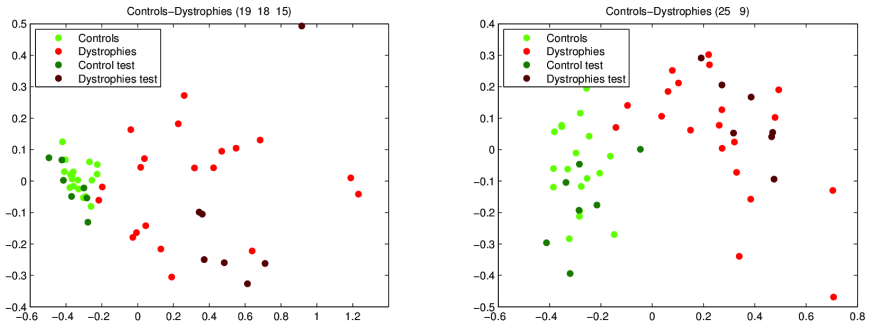
As a summary, it seems coherent to think that the use of structural features improves the discrimination between controls and neurogenic atrophies (see Tables 2.10 and 2.11). However, adding structural features does not improve or even does diminish the classification success rate for the discrimination between controls and dystrofies (C/D) (see Tables 2.10 and 2.11). Therefore, for this case (C/D) a further study of severity grading is proposed.

2.9 Severity grading

The last study involves the degree of severity that a biopsy with dystrophic pattern presents. The affectation degrees of these biopsies were evaluated by a pathologist. A principal component analysis (PCA) is applied to achieve the purpose.

In Fig. 2.12 the two first principal components are represented in a bidimensional space for each of the two sets of features. Green and red points represent control and dystrophy biopsies, respectively. The 7 control biopsies and the 7 dystrophy biopsies used for the evaluation are displayed in dark green and dark red, respectively. The Pearson correlation coefficient is used to study if there is a relationship between the location in the PCA graph and the affectation degree of a biopsy. It is computed between the severity degrees and the Euclidean distances between these pathological images and the centroid of the controls in the PCA graph. This correlation coefficient measures the strength of linear dependence between these two variables (degree and distance). Results are presented in Table 2.14 for the two feature sets. We can observed that the set of 24 features is more correlated with the pathologist's evaluation. And the high values obtained indicate that PCA can be considered as a way to analyse the affectation degree.

Fig. 2.13 shows the PCA graphs for the four remaining comparisons: C/A and C/D/A with each of the two sets of features. Neurogenic atrophies are displayed in blue, the five used as test are represented as light blue points and the atrophy with pseudo-dystrophic nature, also used as test biopsy, in black. It is noteworthy that this last case is located between the two big groups of dystrophies and atrophies (see Fig. 2.13 (d)), since this biopsy presents characteristics of the two diseases, what shows the good classification



(a) (b)

Figure 2.12 PCA graphs.(a) Quadriceps controls and dystrophies with the features 19,18,15 as input (b) Quadriceps controls and dystrophies with features 25, 9 as inputs.

Table 2.14 Pearson Correlation between the affectation degree of the dystrophy biopsies and the Euclidean distance between the pathological images and the centroid of the controls in the PCA graph for the two feature sets.

Comparison	24 features	
	Selected features	Pearson Coeff.
Controls-dystrophies (C/MD)	19 18 15	**0.875**
	82 features	
	Selected features	Pearson Coeff.
Controls-dystrophies (C/MD)	25 9	0.7512

achieved.

Correlation results are not presented for these comparisons because the severity degrees of neurogenic atrophies were unknown.

2.10 Summary and Conclusions

In this chapter a procedure to analyse and classify neuromuscular diseases through biopsy images of fluorescence microscopy is presented. The patterns that muscular biopsies present are dystrophic and atrophic. The procedure begins with an accurate segmentation of the muscle fibres (Section 2.5). Then, two set of features are extracted: 24 features that physicians take into account to diagnose the diseases (Tables 2.5 and 2.6), and 58 structural features that the human eye does not see (Table 2.7), constituting a set of 82 features (24+58). The aim is to study the goodness of these sets and to analyse if the addition of new features improves the classification of the diseases. To developed this study three

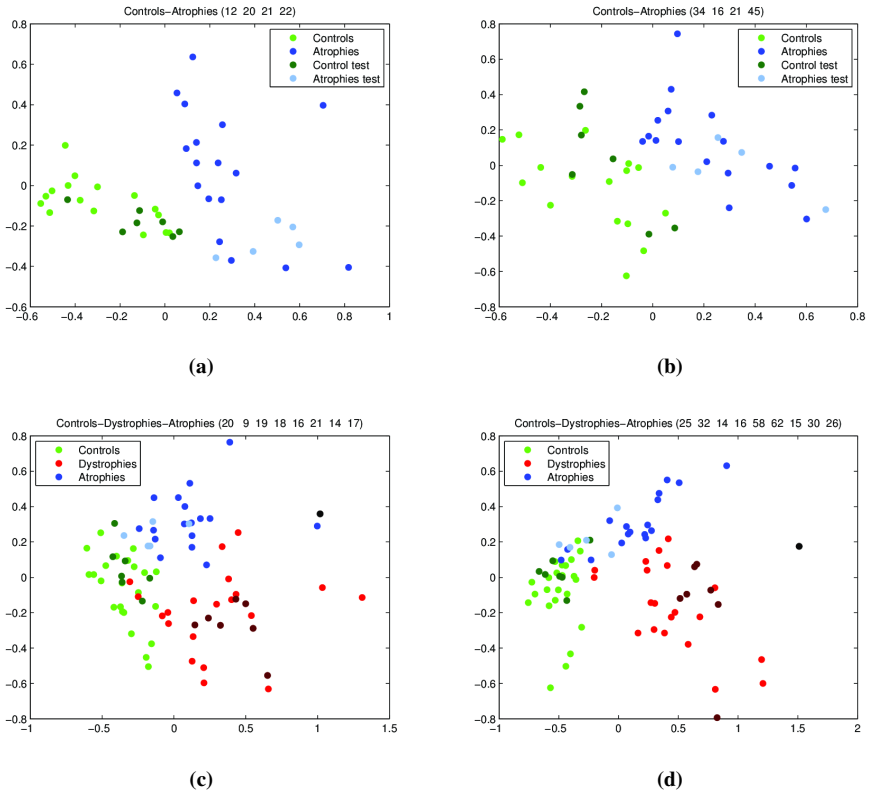


Figure 2.13 PCA graphs.(a) Biceps controls and atrophies with features 12, 20, 21, 22 as input (b) Biceps controls and atrophies with features 34, 16, 21, 45 as inputs (c) Controls, dystrophies and atrophies with the features 20, 9, 19, 18, 16, 21, 14, 17 as inputs (d) Controls, dystrophies and atrophies with features 25, 32, 14, 16, 58, 62, 15, 30, 26 as inputs .

comparisons are presented: control/muscular dystrophies (C/MD), control/neurogenic atrophies (C/NA) and controls/muscular dystrophies/neurogenic atrophies (C/MD/NA). A feature selection step is performed for each comparison and each set of characteristics (Table 2.8 and Table 2.9). 71 training images were used. The classification error obtained for the selected features for the two first comparisons is 0%, however for the triple comparison (C/MD/NA) the classification error decreases with the second set of features. Finally, the classification on 20 test images is carried out by a Fuzzy ARTMAP neural network, whose output is modified to allow the categorization of a new biopsy and the estimation of the certainty of belonging to this category. Results (Table 2.10 and Table 2.11) show that to distinguish between control and atrophies (C/NA) the second set of features improves the classification, however there is no difference for the other two comparisons. In addition, a study to grade the severity of a dystrophic pattern is performed via principal

component analysis for which a correlation coefficient between the grading according to the pathologist and according to PCA has been computed to validate the study (Fig. 2.12 and Table 2.14).

It can be concluded that, in view of all the results presented, the best set of features to distinguish between controls and muscular dystrophies (C/MD) is the first one, derived from the features that the physicians take into account in their diagnosis. This fact may be due to a clear difference between the biopsies belonging to these two groups. However, when biopsies belonging to atrophies come into play (C/NA and C/MD/NA), the addition of new features undetectable by the human visual inspection improves its categorization. This is due to the fact that biopsies affected by atrophy present similar characteristics to control biopsies.

Therefore, the process to classify a new biopsy would be as follows:

- Segmentation of the muscular fibres.
- Extraction of features.
- Classification with the second feature set into three groups: C/MD/NA (features number 25,32,14,16,58,62,15,30 and 26).
- Study of output values (k_{cl}) of the neural network.
- If the output values do not present a clear result, i.e. a k_{cl} significatively higher than the other ones, a classification with the other two comparisons (controls/dystropies and controls/atrophies) is required.
- Application of the criterion shown in Table 2.13 (C=control MD=muscular dystrophy NA=neurogenic atrophy).
- Application of the principal component analysis to estimate the severity degree of the dystrophic patterns.

As future works, a further study of different methods of feature selection, classification and severity grading could be performed, focusing into the categorization of biopsies affected by atrophies because of its similitude with controls.

PART II

Pigmented lesions

CHAPTER 3

Related Publications

- International journals:

  Sáez A., Mendoza C.S., Acha B., Serrano C. Development and evaluation of perceptually adapted colour gradients. *IET Image Processing* 7 (4) , pp. 355-363 (2013).

- International conferences:

  Sáez A., Serrano C., Acha B. Evaluation perceptual color edge detection algorithms. *5th European Conference on Colour in Graphics, Imaging, and Vision and 12th International Symposium on Multispectral Colour Science 2010*, CGIV 2010/MCS'10 , pp. 222-227 (2010).

  Sáez A., Fondoń I., Acha B., Jimeńez S., Alemany P., Abbas Q., Serrano C. Optic disc segmentation based on level-set and colour gradients. *6th European Conference on Colour in Graphics, Imaging, and Vision 2012*, CGIV 2012 , pp. 121-125 (2012).

3 Segmentation based on colour gradients

Todo es según el color del cristal con que se mira.

Ramón de Campoamor

In this chapter an edge-based colour segmentation algorithm is presented. This edge detector is based on a colour gradient and it is implemented by a level set technique. The choice of the colour gradient is derived from a study of a set of colour gradients based on colour visual perception, which use CIE $L^*a^*b^*$ colour space. The main objective of this study is the analysis of how the colour difference equations, developed by International Commission on Illumination (CIE), affect the estimation of the gradients in terms of correlation with colour visual perception. To evaluate the gradients performance they are used as the basis of the edge detector based on level set. A set of synthetic images is designed to evaluate which edge detector and consequently, which colour difference equation, is more correlated with human perception of colour. Both quantitative and qualitative measurements showed that the results obtained using CIE94 colour difference formula have a higher correlation with what the human eye can perceive.

Although this segmentation algorithm for colour images is a general purpose approach, it has been successfully applied in different medical applications, such as optic disc detection in retinography images and pigmented lesion segmentation in dermoscopic images.

3.1 Introduction

Gradient operators are generally defined for greyscale images and mainly used in image processing for edge detection. However, nowadays, many image processing tasks are developed for colour images. The advantage of colour edge detection schemes over greyscale approaches is easily demonstrated by considering the fact that those edges that

exist at the boundary between regions of different colours cannot be detected in greyscale images if there is no change in intensity [Evans and Liu, 2006].

In the recent literature there are many works that deals with colour image gradient research. In [Moreno et al., 2010] an image colour gradient preserving colour constancy is presented. Kyrlov and Nasonov use a colour gradient information to develop an edge-directed image interpolation technique [Krylov and Nasonov, 2011] and in [Romaní et al., 2010] a method for contour detection of cells based on colour gradient is proposed. These works use perceptually non-uniform colour spaces, however some authors seek for the adaptation to the visual colour perception. A colour image segmentation algorithm is proposed in [Garcia Ugarriza et al., 2009], which exploited the information obtained by computing the gradient in CIE $L^*a^*b^*$ colour space. Xue-Wei and Zhang present a perceptual colour edge detection algorithm using CIE CAM02 and iCAM to predict the colour appearance [Xue-Wei and Zhang, 2008]. These authors draw that recent research indicates that most of the colour image segmentation algorithms are very sensitive to colour difference calculation or colour similarity measure [Wesolkowski et al., 2000], [Cheng et al., 2001]. It is safe to say that the accuracy of colour difference calculation determines the performance of various colour image segmentation approaches. Colour difference calculation relies heavily on the uniformity of colour space [Xue-Wei and Zhang, 2008].

One of the approximately uniform colour spaces first recommended by CIE for the purpose of measuring colour differences was CIE $L^*a^*b^*$. CIELAB [McLaren, 1976] is the colour difference equation of CIE $L^*a^*b^*$ colour space. However, it was soon found that it did not fit well to the experimental data with small to medium magnitudes of colour differences. To improve the performance of CIE $L^*a^*b^*$ colour space, a number of datasets were constructed and modified colour difference equations were defined. Some new colour difference equations, such as CIE94 [Report, 1995] and CIEDE2000 [CIE Technical, 2002], were derived.

In the literature the performance of these equations have been evaluated on different data sets [Mandic et al., 2006] [Granger, 2008] [Shen and Berns, 2011] [Wang et al., 2012] and for different applications, above all, for the textile industry [Habekost, 2007] [Aspland and Shanbhag, 2004]. However, the study of how these colour difference equations (CIELAB, CIE94, CIEDE2000) affect in the estimation of the gradient, which is the basis of numerous applications of computer vision and image processing such as segmentation, has not ever been addressed.

In the work proposed, the computation of the gradient is addressed by the Sobel operator generalised to the multidimensional case. This choice was motived by the conclusion of Wesolkowski [Wesolkowski et al., 2000], who drew that the performance of the Sobel operator is superior to others.

To evaluate their performances, the gradients were used as edge detectors. The simplest edge detector is obtained by thresholding those gradients. We compare three cases: gradient calculated with CIELAB, with CIE94 and with CIEDE2000. However, in the three proposed colour gradient estimators, the dynamic range differs between the three colour difference measurements. This implies that a different threshold should be chosen for each gradient. To avoid that the quality of the edges detected by each gradient estimator depends strongly on the choice of this threshold, we propose the use of a level set formulation [Li et al., 2005], in which these colour gradients control external energies.

Following the idea of the evaluation of these gradients based on colour difference equations, we also contribute with a colour image database. It consists on 96 images, generated from the five colour centres recommended by CIE in the guidelines published to coordinate research into colour differences [Robertson, 1978].

Both quantitative and qualitative measurements were used to test the correlation between the visually perceived colour difference and the detector output, and consequently, the colour difference equations.

The edge detector with the best performance will be applied as segmentation algorithm for pigmented lesions in dermoscopic images.

The rest of the chapter is organized as follows: Section 3.2 summarizes the colour gradients proposed. The implemented variational level set technique is introduced in Section 3.3. In Section 3.4 the complete developed method is exposed. The experimental results are explained in Section 3.5. Its application to segment pigmented lesions is presented in Section 3.6. Some conclusions are exposed in Section 3.7.

3.2 Colour gradients

Several perceptual colour gradients are proposed in this section. They are developed in a uniform colour space and they preserve the vectorial nature of the colour images in that space. Before their explanation, we introduce the uniform colour space and the colour difference equations, in which they are based on.

3.2.1 Perceptual Uniform Colour Space and Colour Difference Equations

In 1976, the International Commission on Illumination (CIE) standardized $L^*a^*b^*$ colour representation system as perceptually uniform [McLaren, 1976]. A colour space is perceptually uniform if perceptual colour differences can be measured with Euclidean distances in that space. The three coordinates of $L^*a^*b^*$ represent the lightness of the colour (L^*), its position between red/magenta and green (a^*) and its position between yellow and blue (b^*). This can also be expressed in terms of cylindrical coordinates with the perceived lightness L^*, the chroma C^*_{ab} and the hue h^*_{ab}, defined in (3.1) and (3.2) respectively.

$$C_{ab}{}^* = \sqrt{a^{*2} + b^{*2}} \tag{3.1}$$

$$h_{ab} = arctan\left(\frac{b^*}{a^*}\right) \tag{3.2}$$

The CIELAB colour difference, also known as ΔE^*_{ab}, is calculated using (3.3).

$$\Delta E^*_{ab} = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \tag{3.3}$$

where $\Delta L^* = L^*_1 - L^*_2$ (Δa^* and Δb^* are defined in the same manner for coordinates a^* and b^*).

However, subsequent experiments demonstrated that Euclidean distance ΔE^*_{ab} is not an accurate measurement of perceived colour difference between two stimuli. To correct the

problem a new difference formula was recommended by CIE [Report, 1995] in 1994.

$$\Delta E_{94}{}^{*} = \sqrt{\left(\frac{\Delta L^{*}}{k_L S_L}\right)^2 + \left(\frac{\Delta C^{*}_{ab}}{k_C S_C}\right)^2 + \left(\frac{\Delta H^{*}_{ab}}{k_H S_H}\right)^2} \tag{3.4}$$

$$S_L = 1, \tag{3.5}$$

$$S_C = 1 + 0.045\, C^{*}_{ab}, \tag{3.6}$$

$$S_H = 1 + 0.015\, C^{*}_{ab}. \tag{3.7}$$

The factors k_L, k_C and k_H, are included to match the perception of the background conditions.

Later, CIEDE2000 was developed to correct deficiencies in previous colour difference equations [Luo et al., 2001]. Its accuracy in predicting small perceived colour differences has been demonstrated [Cui and Luo, 2009].

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{K_L S_L}\right)^2 + \left(\frac{\Delta C'}{K_C S_C}\right)^2 + \left(\frac{\Delta H'}{K_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{K_C R_C}\right)\left(\frac{\Delta H'}{K_H S_H}\right)} \tag{3.8}$$

where $\Delta C'$ and $\Delta H'$ are defined in [CIE Technical, 2002].

3.2.2 Proposed colour gradients

In colour gradients, the vectorial nature of colour is preserved throughout the computation. Colour images are viewed as a two-dimensional three channel vector field. Each channel in this vector is characterized by a discrete integer function $\mathbf{f}(x,y)$. The value of this function at each point is defined by a three dimensional vector in a given colour space [Rangayyan et al., 2011]. Therefore, a pixel is defined as in (3.9).

$$\mathbf{f}(x,y) = \begin{bmatrix} C_1(x,y) \\ C_2(x,y) \\ C_3(x,y) \end{bmatrix} \tag{3.9}$$

where $C_i(x,y)$ represents the value of the pixel in the i-th colour plane ($i = 1,2,3$), and (x,y) refers to the spatial dimensions in the 2-D plane.

$\mathbf{f}_9$	$\mathbf{f}_5$	$\mathbf{f}_7$
$\mathbf{f}_2$	$\mathbf{f}_1$	$\mathbf{f}_3$
$\mathbf{f}_6$	$\mathbf{f}_4$	$\mathbf{f}_8$

Figure 3.1 Sliding window.

$$X_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, X_2 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{3.10}$$

The operator based on the first derivative, commonly applied in greyscale imaging, is generalized into the multidimensional case in [Plataniotis and Venetsanopoulos, 2000]. Plataniotis extends the Sobel operator (see the horizontal and vertical masks in (3.10)), by constructing the following vectors:

$$\mathbf{H}^+ = \mathbf{f}_7 + 2\mathbf{f}_3 + \mathbf{f}_8 \tag{3.11}$$
$$\mathbf{H}^- = \mathbf{f}_9 + 2\mathbf{f}_2 + \mathbf{f}_6 \tag{3.12}$$
$$\mathbf{V}^+ = \mathbf{f}_6 + 2\mathbf{f}_4 + \mathbf{f}_8 \tag{3.13}$$
$$\mathbf{V}^- = \mathbf{f}_9 + 2\mathbf{f}_5 + \mathbf{f}_7 \tag{3.14}$$

according to the notation used in Fig. 3.1. The colour vector gradients were calculated as $\mathbf{H}^+ - \mathbf{H}^-$ and $\mathbf{V}^+ - \mathbf{V}^-$, respectively. To estimate the colour variation in the vertical and horizontal axes, the following scalars were calculated: $\|\mathbf{H}^+ - \mathbf{H}^-\|$, $\|\mathbf{V}^+ - \mathbf{V}^-\|$. The magnitude B of the maximum variation was estimated as: $B = \sqrt{\|\mathbf{H}^+ - \mathbf{H}^-\|^2 + \|\mathbf{V}^+ - \mathbf{V}^-\|^2}$.

However, we propose calculating the gradient along the x and y axes as shown in (3.15) and (3.16).

$$G_x = \Delta E\left(\mathbf{H}^+, \mathbf{H}^-\right) \tag{3.15}$$

$$G_y = \Delta E\left(\mathbf{V}^+, \mathbf{V}^-\right) \tag{3.16}$$

where ΔE denotes the colour difference between two vectors. In this thesis three colour difference distances are studied, obtaining, consequently, three colour gradients by applying:

- $\mathbf{f}(x,y) = [L^*(x,y), a^*(x,y), b^*(x,y)]$ and ΔE determined by the Euclidean distance (CIELAB).
- $\mathbf{f}(x,y) = [L^*(x,y), a^*(x,y), b^*(x,y)]$ and ΔE determined by CIE94 colour difference equation.
- $\mathbf{f}(x,y) = [L^*(x,y), a^*(x,y), b^*(x,y)]$ and ΔE determined by CIEDE2000 colour difference equation.

The gradient magnitude can then be computed as shown in (3.17).

$$VD = \sqrt{G_x^2 + G_y^2} \tag{3.17}$$

3.3 Variational level set

Once a colour gradient is estimated, the simplest edge detector is obtained by thresholding that gradient. In the three proposed colour gradient estimators, the dynamic range of the colour gradient differs between the three colour difference measurements. This implies that a different threshold should be chosen to detect edges with each different gradient estimator and the quality of the edges detected from each gradient estimator depends strongly on the choice of this threshold. This is the reason why a level set formulation was applied as edge detector, and thus, to be able to compare the three gradient estimators regardless of the threshold choice .

The Level set method for capturing dynamic interfaces and shapes was introduced by [Osher and Sethian, 1988]. The basic idea is to represent contours as the zero level set of a higher dimensional function, usually referred as the level set function (LSF)(ϕ), and to formulate the motion of the contour as the evolution of the level set function. Since this method was proposed, it has had far-reaching impact in various applications, such as computational geometry, fluid dynamics, image processing and computer vision. However, in image processing and computer vision applications, the level set method was introduced independently by [Caselles et al., 1994] and [Malladi et al., 1995a] in the context of active contour models [Kass et al., 1988] for image segmentation. Early active contour models are formulated in terms of a dynamic parametric contour $\mathbf{C}(s,t)$. The curve evolution can be expressed as:

$$\frac{\partial \mathbf{C}(s,t)}{\partial t} = F\mathbf{N} \tag{3.18}$$

where F is the speed function that controls the motion of the contour, and $\mathbf{N}$ is the inward normal vector to the curve C. The curve evolution in this equation can be converted to a level set function by embedding the dynamic contour $C(s,t)$ as the zero level set of a time dependent level set function $\phi(x,y,t)$ with spatial variables x,y and a temporal variable t. Then, the curve evolution equation (equation (3.18)) is converted to a partial differential equation (PDE):

$$\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0 \tag{3.19}$$

which is called level set evolution equation. This approach presents the advantage that the level set function may break or merge naturally during the evolution. The evolution PDE of the level set function can be directly derived from the problem of minimizing a certain energy functional defined on the level set function. This type of methods are known as variational level set methods and they are more convenient for incorporating additional information, such as region-based information [Chan and Vese, 2001] and shape-prior information [Malladi et al., 1995b].

In this thesis, the variational level set formulation proposed by [Li et al., 2005] [Li et al., 2010], which includes a distance regularization term and an external energy term that drives the motion of the zero level contour toward desired locations, is used. The use of this distance regularization term is motived because in conventional level set methods, the LSF typically develops irregularities during its evolution, which cause numerical errors

and eventually destroy the stability of the level set evolution. To overcome this difficulty, a numerical remedy, commonly known as reinitialization [Sethian, 1999], was introduced to restore the regularity of the LSF and maintain stable level set evolution. Reinitialization is performed by periodically stopping the evolution and reshaping the degraded LSF as a signed distance function. Li et al. [Li et al., 2010] avoided this reinitialization with the inclusion of the distance regularization term, defined with a potential function, which maintains a desired shape of the level set function, particularly a signed distance profile near its zero level set (for the 2-D case the signed distance function is considered as a surface $z = \phi(x,y)$).

In the reported works [Li et al., 2005] and [Li et al., 2010], it was shown that the PDE in (3.19) can be expressed as:

$$\frac{\partial \phi}{\partial t} = \mu \left[\nabla^2 \phi - div\left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] - \frac{\partial \varepsilon_{ext}}{\partial \phi} \tag{3.20}$$

where the first right hand term is associated with the distance regularization and the second with the external energy. This formulation can be used in different applications with different definitions of the external energy (ε_{ext}). Since our proposal is image segmentation, we use edge-based information in the external energy. Li defined an edge indicator function by:

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^2} \tag{3.21}$$

where G_σ is the Gaussian kernel with standard deviation σ and I is the test image. The convolution in 3.21 is used to smooth the image to reduce the noise. This function usually takes smaller values at object boundaries than at other locations.

For a level set function $\phi : \Omega \to \Re$, the authors defined an external energy functional $\varepsilon_{ext}(\phi)$ by:

$$\varepsilon_{ext}(\phi) = \lambda L_g(\phi) + \alpha A_g(\phi) \tag{3.22}$$

where $\lambda > 0$ and $\alpha \in \Re$ are coefficients of the energy functional $L_g(\phi)$ and $A_g(\phi)$. The energy $L_g(\phi)$ is minimized when the zero level contour of ϕ is located at the object boundaries. The energy $A_g(\phi)$ is introduced to speed up the motion of the zero level contour in the level set evolution process, which is necessary when the initial contour (ϕ_0) is placed far away from the desired object boundaries. These energies are defined by:

$$L_g(\phi) = \int_\Omega g\delta(\phi)|\nabla \phi| dxdy \tag{3.23}$$

$$A_g(\phi) = \int_\Omega gH(-\phi)dxdy \tag{3.24}$$

where δ and H are the Dirac delta function and the Heaviside function, respectively. In practice, the Dirac delta function and Heaviside function in the functionals are approximated by the following smooth functions and as in many level set methods [Osher and

Fedkiw, 2003], defined by:

$$\delta_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon}\left[1+cos\left(\frac{\pi x}{\varepsilon}\right)\right], & |x| \leq \varepsilon \\ 0, & |x| > \varepsilon \end{cases} \tag{3.25}$$

$$H_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon}\left(1+\frac{\pi}{\varepsilon}+\frac{1}{\pi}sin\left(\frac{\pi x}{\varepsilon}\right)\right], & |x| \leq \varepsilon \\ 1, & |x| > \varepsilon \\ 0, & |x| < -\varepsilon \end{cases} \tag{3.26}$$

The parameter ε is usually set to 1.5.

It is shown in [Li et al., 2005] and [Li et al., 2010] that the energy functional can be minimized by solving the following gradient flow:

$$\frac{\partial\phi}{\partial t} = \mu\left[\Delta\phi - div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right] + \lambda\,\delta_\varepsilon(\phi)div\left(g\frac{\nabla\phi}{|\nabla\phi|}\right) + \alpha g\delta_\varepsilon(\phi) \tag{3.27}$$

given an initial LSF $\phi(x,0) = \phi_0(x)$. Regarding the time step (τ), the authors found that τ and the coefficient μ must satisfy $\tau\mu < \frac{1}{4}$ in order to maintain stable level set evolution. Using larger time step can speed up the evolution, but may cause error in the boundary location if the time step is chosen too large. There is a trade-off between choosing larger time step and accuracy in boundary location. A $\tau \leq 10$ is used for most of the images.

In our proposal, the edge indication function g is modified to:

$$g = \frac{1}{1+|VD(diff\{I\})|^2} \tag{3.28}$$

where $diff$ is an anisotropic diffusion filter and VD are the proposed colour gradients that are applied to the diffused image. Modifying the edge indicator function involves two aspects:

- The Gaussian filter smoothing is substituted by a colour anisotropic diffusion filter [Lucchese and Mitra, 2001]. Both forms of smoothing were tested but the colour anisotropic diffusion filter obtained better results in the experiments.
- Gradients in the diffused image were computed with the three proposed colour gradients approaches explained in Section 3.2.2.

3.4 Methodology

The flow diagram in Fig. 3.2 gives an overview of the main steps of the segmentation algorithm implemented, that moreover allow us to evaluate colour gradients.

a. Uniform colour space transform. The RGB image is transformed into the uniform colour space CIE $L^*a^*b^*$, described previously in Section 3.2.1.

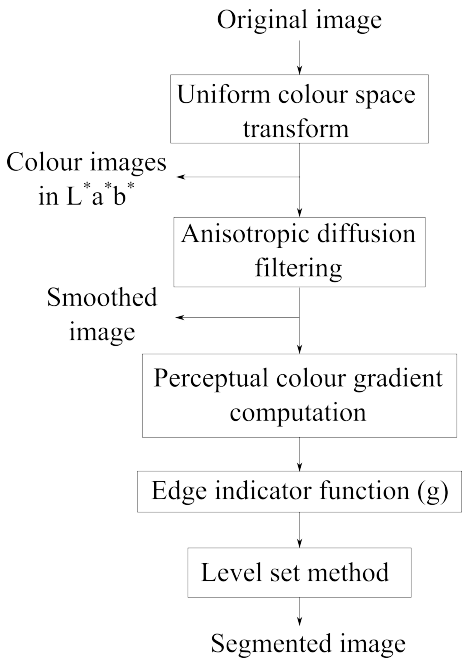


Figure 3.2 Proposed segmentation system.

b. Anisotropic diffusion filtering. In the anisotropic diffusion filtering [Perona and Malik, 1990], within-region smoothing is performed largely without blurring region boundaries. In this thesis an extension of the method to colour images is implemented [Lucchese and Mitra, 2001], diffusing separately the chromatic and achromatic channels. This way has its rationale in colour vision models: the human visual system senses colour information through photoreceptors which can be regarded as three sets of filters tuned to the wavelengths of red, green and blue; this information is then split into chromatic (2-D) and achromatic (1 -D) channels before being further and independently processed. Neurophysiological evidence shows that there exists a perfect agreement between the second Human Visual System processing stage and the opponent-colours theory based on the three antagonistic mechanisms red-green, blue-yellow and black-white [Rangayyan et al., 2011]. These stimuli can also be conveniently expressed in terms of hue, saturation and lightness. Hue and saturation are processed together under the name of phasors: hue is the phase and saturation is the magnitude of a complex function defined as complex chromaticity. The scalar achromatic information represented by lightness is diffused separately.

c. Perceptual colour gradients computation. The colour gradients proposed in Section 3.2.2 are computed from the diffused image.

d. Level set technique. The level set formulation explained in Section 3.3 is applied

with the edge indicator function defined in 3.28.

3.5 Results

An extensive comparative study between the results of the detectors based on the three colour difference equations was carried out to test which of them best correlates to colour visual perception.
A 96 synthetic image database was generated to evaluate the performance of the gradients. To this end, evaluation procedures defined by [Zhu et al., 1999] were carried out. Both quantitative and qualitative measurements were used. The quantitative performance measurements were based on edge deviation from true edges. For this experiment the predefined edge map (ground truth) was required. Since objective measurements are not sufficient to model the complexity of human visual systems, a qualitative evaluation was also performed. This evaluation made it possible to determine the correlation between the detector output and the visually perceived colour differences.

3.5.1 Images database

A set of 96 colour images was created following CIE guidelines to coordinate research into colour differences [Robertson, 1978]. Five colour centres were recommended for study. Our images are based on these centres, which are shown in Table 3.1.

Table 3.1 Colour centres following CIE guidelines, given in L^*, a^*, b^*, C^* and h coordinates .

Colour	L^*	a^*	b^*	C^*	h
Grey	63.5	-0.6	0.8	0.9	126.4
Red	46.2	37.8	23.8	44.7	32.2
Yellow	87.9	-6.6	46.1	46.5	98.2
Green	58.6	-33.7	0.8	33.7	178.7
Blue	37.3	4.7	-32	32.3	278.3

In each image of the database two colour centres from Table 3.1 are present. Along with these two colour centres, two additional colours appears in the image. These two additional colours are X CIELAB units distant from those colour centres. X is a parameter that varies along the database.

To evaluate the ability to detect different edge orientation, rectangular and circular objects were included in the different images of the database. Fig. 3.3 (a),(d),(e) and (f) shows example images from the dataset. In Fig. 3.3 (a) there is a small square inside the green zone with 0.5 CIELAB units colour difference to the green background (green centre), and a small square inside the yellow square with 0.5 CIELAB units to the yellow background (yellow centre) as its ground-truth map indicates (Fig. 3.3 (b)). However,

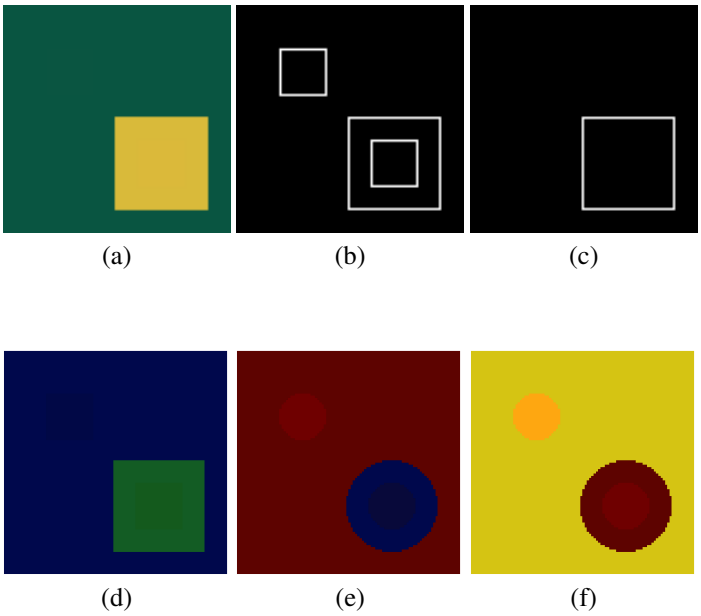


Figure 3.3 Example of images from the dataset. (a) Sample pairs with 0.5 CIELAB units. (b) Ground-truth map of (a). (c) Perceived ground-truth map of (a). (d) Sample pairs with 4 CIELAB units. (e) Sample pairs with 10 CIELAB units. (f) Sample pairs with 10 CIELAB units.

the human visual system cannot perceive this difference, therefore, the perceived ground-truth map of this image would be as shown in Fig. 3.3 (c). Despite the fact that the colour difference in Fig. 3.3 (b) is greater (4 CIELAB value) the small squares are still almost imperceptible. Nevertheless in Fig. 3.3 (c) and (d) red, blue and yellow samples can now be perceived by the human visual system with 10 CIELAB units between samples. It would be desirable to develop an edge detector that find the boundaries that human visual system is able to detect.

It is also interesting to note that for two sample pairs with the same CIELAB units between them, visual assessment may not be equal. An example of this effect is illustrated in Fig. 3.4, where the difference in terms of CIELAB units between grey samples is the same as between the blue samples, but the perceptual difference is not the same. We perceive more difference between the grey pair than the blue pair. However, if this perceptual difference is represented in CIE94 and CIEDE2000 equations, as shown in Table 3.2, values in CIE94 units and CIEDE2000 units are lower for the blue pair than for the grey pair as the human eye perceives them.

Thus, the aim is to study how the colour difference equations influence the estimation of the colour gradients and check which is more capable of behaving like the human eye.

Since the method uses a level set formulation, a initial curve (ϕ_0) is required to begin the process (see Section 3.3). To evaluate the results with the proposed database, two

Table 3.2 The CIELAB values of colour centres.

	CIELAB	CIE94	CIEDE2000
Grey	10	9.613	10.259
Blue	10	4.17	5.6

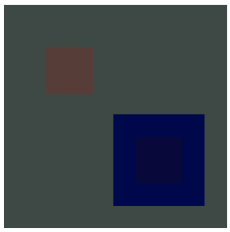


Figure 3.4 Example of images with same CIELAB units between two pairs but different visual perception .

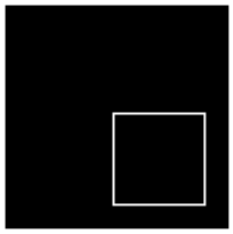


Figure 3.5 Initial curve (ϕ_0) for level set formulation.

initial curves are used. One is the interior edge of the complete image and the second the interior edge of the smaller square or circle as shown Fig. 3.5.

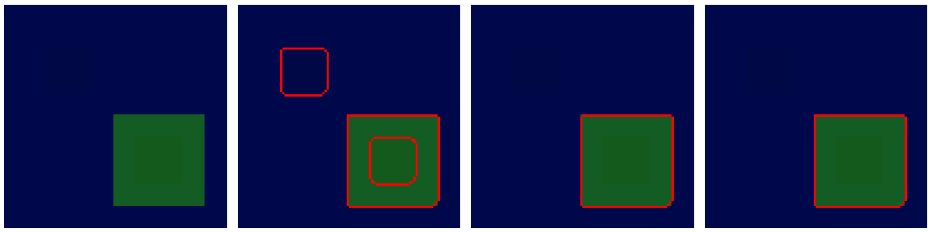
3.5.2 Evaluation of results

Two experiments were carried out to evaluate the resulting edge maps: a qualitative measurement through a subjective test and a quantitative evaluation measurement. The quantitative measurement evaluated how much the detected edge deviates from the true edges (since images are synthetic we know the ground truth). The subjective test provided information about how much the edge detection correlates with the visually perceived colour differences.

3.5.2.1 Qualitative measurement: Subjective test

Subjective evaluation is very important in image processing [Plataniotis and Venetsanopoulos, 2000]. Moreover, as the aim of our work is related to visually perceived colour differences, this issue becomes essential.

In this evaluation, a panel of 6 observers with normal colour vision, who passed the Ishihara test [Ishihara, 1998], were asked how many different colours they could distinguish in each image. For each image, the true number of perceived colours was fixed to the most voted among the observers. There was a low deviation in the answers of the observers; 68 answers of the 96 were unanimously chosen, 16 with 5 votes and 12 with 4 votes. A detector out got one hit in an image when the number of colours detected in that image coincide with the true number of perceived colours. For example, in Fig. 3.6 (a), the observers distinguished two different colours. The detectors based on CIE94 and CIEDE2000 (Fig. 3.6 (c) and (d)) detected two different colours, but the detector based on Euclidean distance distinguished four colours (Fig. 3.6 (b)) so, CIE94 and CIEDE2000 detectors got one hit each of them and CIELAB did not get any. The average hit ratio for each detector is presented in Table 3.3.



(a) (b) (c) (d)

Figure 3.6 (a) Test image. (b) Output of detector based on CIELAB. (c) Output of detector based on CIE94. (d) Output of detector on CIEDE2000.

Table 3.3 Subjective test results.

	Lab	CIE94	CIEDE2000
Hit Ratio	65.54%	80.2%	76.04%

As shown in Table 3.3, the performance of the detector based on CIE94 correlates more closely to human colour perception. This result contrasts with the fact that CIEDE2000 was developed to improve CIE94.

3.5.2.2 Quantitative measurement

An edge map is defined as a binary image where edges are valued 1. The ground-truth edge map is the edge map which contains the real edges. Quantitative evaluation requires the ground-truth edge map. In the proposed evaluation we have distinguished between two ground-truth edge maps: the objective ground-truth edge map, known because the database contains only synthetic images, and the perceived ground-truth edge map, containing edges according to observers' perception. The objective ground-truth edge map

and the perceived ground-truth edge map differ in some images in the database. This is illustrated in Fig. 3.6. In Fig. 3.6 (a) four colours are present but only two are perceived.

The first quantitative measurement used to evaluate the three colour detectors is the figure of merit proposed by Pratt [Pratt, 1978]. It is defined as:

$$F_{Pratt} = \frac{\Sigma_{i=1}^{I_D} \left(\frac{1}{1+\alpha d(i)^2} \right)}{max(I_D, I_I)} \tag{3.29}$$

where I_D is the amount of pixels that the detector considers edges, I_I is the amount of real pixels belonging to an edge, $d(i)$ is the distance between the i-th pixel of the detector edge map and its nearest pixel in the real edge map, and α is a scaling constant with a usual value of 1/9. When F_{Pratt} is 1, the computed edge matches the real edge.

Moreover, we adopted another method to evaluate the performance of edge detectors [Boaventure I., 2009]. The proposed index is a combination of the figure of merit F_{Pratt}, the percentage of pixels that were correctly detected (P_{co}), the percentage of pixels that were not detected (P_{nd}) and the percentage of pixels that were erroneously detected as edge pixels (P_{fa}). It is defined by Euclidean distance ($d_{£2}^4$) to the point $P = (1,1,0,0)$, where its coordinates are optimum values achieved by indices P_{co}, F_{Pratt}, P_{nd} and P_{fa} respectively. The point P represents the optimum point to be reached by an ideal edge detector. The distance to this point can be calculated by the equation:

$$d_{£2}^4 = \sqrt{(P_{co}-1)^2 + (F_{Pratt}-1)^2 + P_{nd}^2 + P_{fa}^2} \tag{3.30}$$

where P_{co}, P_{nd} and P_{fa} are defined as:

$$P_{co} = \frac{TP}{max(I_D, I_I)} \tag{3.31}$$

$$P_{nd} = \frac{FN}{max(I_D, I_I)} \tag{3.32}$$

$$P_{fa} = \frac{FP}{max(I_D, I_I)} \tag{3.33}$$

TP (true positive) is the number of correctly detected edge pixels, FN (false negative) is the amount of edge pixels that were not classified as edge pixels and that actually were edge pixels, and FP (false positive) is the number of pixels erroneously classified as edge pixels.

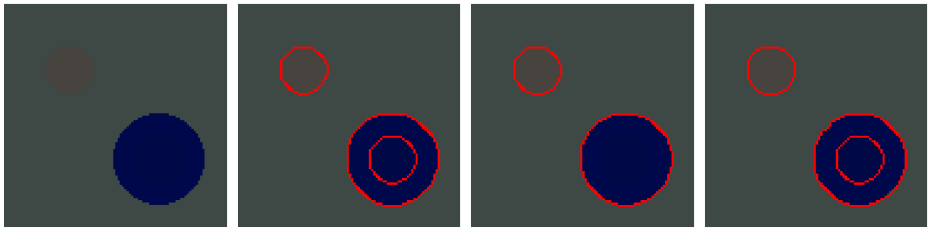
The values of these statistical indices range between 0 and 1, and reach ideal values in case 1 for P_{co} and 0 for indices P_{nd} and P_{fa}. Therefore, the distance defined $d_{£2}^4$ varies between 0 and 2, where the value 0 represents a perfect fit between two edge maps.

As explained above, two measurements were derived from (3.29) and (3.30). In the objective F_{Pratt} and $d_{£2}^4$, the objective ground-truth edge map was used for the evaluation and in the perceived F_{Pratt} and $d_{£2}^4$ edges contained in the perceived ground-truth edge map were used to calculate equation (3.29) and (3.30). The results of the evaluation of F_{Pratt}

and $d^4_{£2}$ with both ground-truth edge maps are presented in Table 3.4. As can be observed, the best result for the perceived F_{Pratt} and $d^4_{£2}$ corresponds to the detector using the CIE94 difference equation, as in the subjective test. This is not the case for the objective F_{Pratt} and $d^4_{£2}$, where the detector based on CIELAB attained the best result.

Table 3.4 F_{Pratt} and $d^4_{£2}$ values.

	CIELAB		CIE94		CIEDE2000	
	Mean	Var.	Mean	Var.	Mean	Var.
Objective F_{Pratt}	0.89	0.1361	0.7753	0.1993	0.7688	0.1926
Perceived F_{Pratt}	0.8404	0.1651	**0.8832**	0.1081	0.869	0.1091
Objective $d^4_{£2}$	0.978	0.151	1.001	0.166	1.012	0.155
Perceived $d^4_{£2}$	0.9488	0.144	**0.8855**	0.096	0.924	0.122



(a) (b) (c) (d)

Figure 3.7 (a) Test image. (b) Output of detector based on CIELAB. (c) Output of detector based on CIE94. (d) Output of detector based on CIEDE2000 .

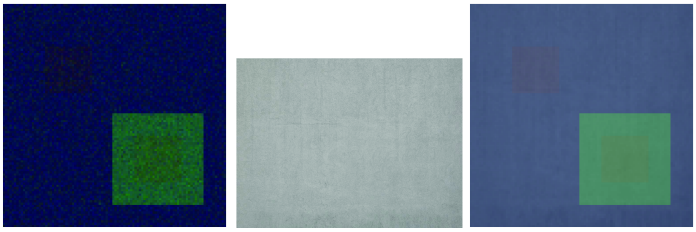
Fig. 3.7 shows another example of the detectors' performance that illustrates the advantages of the detector based on CIE94 colour difference equations. Observers agreed that they could distinguish a blue circle and a small greyish circle on the grey background but they could not distinguish a smaller circle inside the blue circle. The detector based on CIE94 was the only one which detected the circles according to the observers perception.

3.5.2.3 Analysis of noised and textured images

Although the main aim of the study was focused on colour perception, since the final objective is the application to real colour images, two tests were performed regarding to performance of edge detectors in presence of noise and texture.

In the first test, Gaussian noise of zero mean and variance 0.01 is added to the images (see an example in Fig. 3.8 (a)), and in the second one, a texture image (Fig. 3.8 (b)) is added to the images (see an example in Fig. 3.8 (c)).

The same evaluation procedures (subjective and quantitative) have been performed. Table 3.5 shows the hit ratio regarding to the observers perception.



(a) (b) (c)

Figure 3.8 (a) Image from the database in presence of Gaussian noise of zero mean and variance 0.01. (b) Textured image. (c) Image from the database when the textured image (b) is added.

Table 3.5 Subjective test results of the noised and textured images.

	Lab	CIE94	CIEDE2000
Hit Ratio Gaussian noise	43.75%	**70.8**%	60.42%
Hit Ratio Texture	62.5%	**66.67**%	60.42%

The quantitative results are presented in Table 3.6 and Table 3.7

Table 3.6 F_{Pratt} and $d^4_{\pounds 2}$ values for the images in presence of Gaussian noise.

Gaussian noise	CIELAB		CIE94		CIEDE2000	
	Mean	Var.	Mean	Var.	Mean	Var.
Objective F_{Pratt}	0.625	0.2107	0.6142	0.2162	0.621	0.1703
Perceived F_{Pratt}	0.535	0.2234	**0.7184**	0.2145	0.717	0.2119
Objective $d^4_{\pounds 2}$	1.3404	0.2808	1.2516	0.2399	1.2936	0.2216
Perceived $d^4_{\pounds 2}$	1.3527	0.2434	**1.1928**	0.2425	1.2391	0.2623

Although the performance of the detectors diminishes with the noise and textured images, they continue showing good performances and high correlation with human vision and the detector based on CIE94 still gets the results more correlated with the perceived evaluation.

3.6 Segmentation of pigmented lesions

In the last decades, numerous segmentation techniques have been proposed to facilitate the diagnosis of dermoscopy images. The segmentation stage is one of the most important

Table 3.7 F_{Pratt} and $d^4_{\pounds 2}$ values for the textured images.

Texture image	CIELAB		CIE94		CIEDE2000	
	Mean	Var.	Mean	Var.	Mean	Var.
Objective F_{Pratt}	0.6131	0.189	0.6289	0.1933	0.6624	0.2014
Perceived F_{Pratt}	0.7375	0.1712	**0.8355**	0.1322	0.8297	0.1366
Objective $d^4_{\pounds 2}$	1.1906	0.1928	1.1488	0.1767	1.041	0.1628
Perceived $d^4_{\pounds 2}$	1.1367	0.2030	1.0300	0.188	**0.9179**	0.1047

ones for two main reasons [Emre Celebi et al., 2013]. First, the border structure provides important information for accurate diagnosis, as many clinical features such as asymmetry, border irregularity, and abrupt border cutoff are calculated directly from the border. Second, the extraction of other important clinical features such as dermoscopic pattern within lesions, critically depends on the accuracy of border detection.

There have been a great number of algorithms for dermoscopy image segmentation in the literature, to mention some, it can be found methods such as fuzzy c means clustering [Schmid, 1999], thresholding [Ganster et al., 2001] [Yüksel and Borlu, 2009], gradient vector flow (GVF) [Zhou et al., 2011], level set algorithms [Nourmohamadi and Pourghassem, 2012], j-image segmentation algorithm [Emre Celebi et al., 2007], statistical region merging [Celebi et al., 2008b], wavelet transform [Castillejos et al., 2012], wavelet Networks [Sadri et al., 2013].

However, to the best our knowledge, there is no technique superior to the rest. Pigmented lesions segmentation is difficult because of the great variety of lesion shapes, sizes, and colours along with different skin types and textures. In addition, some lesions have irregular boundaries and in some cases there is a smooth transition between the lesion and the skin. Other difficulties are related to the presence of different artefacts, such us dark hair, oil bubbles, black frames or grid markers.

The main advantages in the proposed segmentation method are the application of the procedure without requiring a preprocessing step to remove artefacts, the inclusion of colour information and its automatic nature. The last aspect becomes crucial if the final objective is the development of a CAD system.

Any edge-based level set technique requires an initial curve to begin the process, that can be introduced by the user or can be automatically found. In this work we opt for the second scenario. The steps listed below are carried out to automatically find the initial contour. In Fig. 3.9 an example of each of them is shown.

a) First, the original image is smoothed with a 20×20 spatial averaging filter for multidimensional images and then, a Principal Component Analysis (PCA) by lumping the three channels R, G and B, is applied to this smoothed image in RGB.

b) An Otsu's thresholding is applied to the first principal component image (Fig. 3.9 (b)), resulting a binary image in which, apart from the lesion, artefacts can appear, such as hair or grid markers (see Fig. 3.9 (c)).

c) Finally, artefacts are removed by by shape conditions. As a lesion is supposed to approach a circle, the region of interest corresponds to the one with the biggest area and the lowest eccentricity. The eccentricity is defined as the ratio of the distance between the foci of the ellipse that has the same second-moments as the region and its major axis length. The resulting region is dilated with a disk-shaped structuring element with radius of 5 pixels (Fig. 3.9 (d)) to ensure that the initial contour surrounds the lesion. The contour of this dilated image is the initial contour.

The contour of the resulting image (Fig. 3.9 (d)) is the initial contour.

Once the initial contour is found, the proposed segmentation method based on CIE94 colour gradient is applied, since this gradient was found to have the best performance in the evaluation study. After applying (3.28) in order to compute the edge function, g, this resulting edge map is enhanced by an histogram equalization follow by a lineal expansion of its dynamic range to [0 1] (see Fig. 3.9 (e)). The required constants in the level set formulation are experimentally fixed to $\mu = 0.04$, $\lambda = 9$ and $\alpha = 1.5$. In most of the level set schemes, the curve evolution stops when the iterations reach a fixed number. However, in this work, in an attempt to reduce the computational cost, we propose a stop condition. The curve stops if it does not evolve in two consecutive iterations, what implies that it has reached an object boundary. The time step between each iteration is fixed to 5 so that ($\tau\mu \leq \frac{1}{4}$) for stability in the curve evolution [Li et al., 2005] [Li et al., 2010]. It is important to note that in spite of the presence of artefacts a good segmentation is achieved (Fig. 3.9 (e)).

In Fig. 3.10 and Fig. 3.11 more examples of segmented pigmented lesions are shown.

3.7 Summary and Conclusions

In this chapter a segmentation method has been proposed. Although it is a general purpose framework, results when it is applied to pigmented lesion in dermoscopic images have been presented.

The proposed segmentation method is based on a perceptually adapted gradient integrated in a level-set framework for edge detection in colour images. A previous analysis of three different perceptually adapted colour gradients have been proposed. They use a uniform colour space ($CIEL^*a^*b^*$) and they are based on Euclidean (CIELAB), CIE94 and CIEDE2000 colour differences, respectively. The colour edge detectors derived from these three gradients were evaluated with two tests: a subjective test and quantitative evaluation measurements. To this purpose, a synthetic image database following CIE guidelines for coordinated research on colour difference evaluation [Robertson, 1978] was developed. The edge detector based on CIE94 proved to be the best according to both the subjective test and the perceived quantitative measurements (perceived F_{Pratt} and $d^4_{\pounds 2}$). Another advantage of CIE94 colour difference equation is its low computational cost when compared to CIEDE2000 difference equation. The computation cost of computing the three colour gradients per image on average is presented in Table 3.8. In [Fairchild, 2005] the author states that the complexity of the CIEDE2000 exceeds that of CIE94, as well as, that this high cost is not justified for most practical applications.

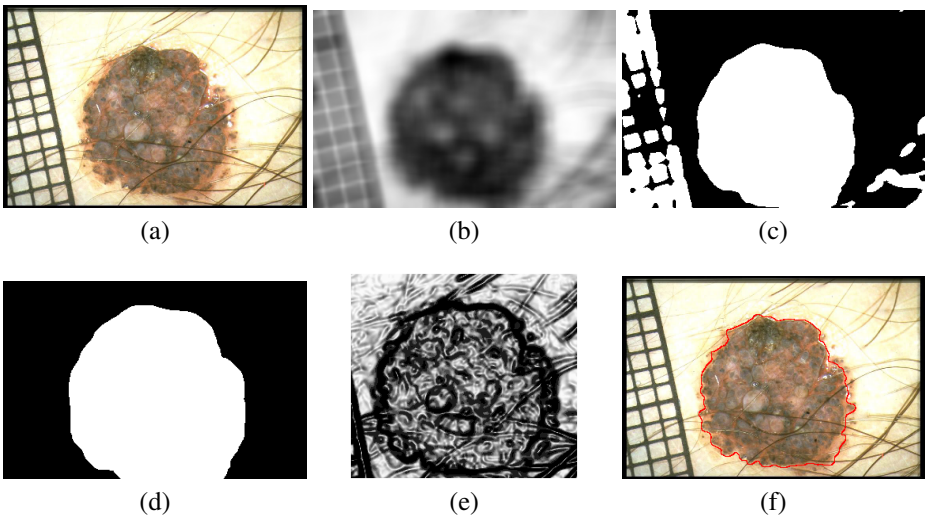


Figure 3.9 Steps followed in the segmentation process. (a) Original image with artefacts: hair and grid marker. (b) First principal component the smoothed RGB image. (c) Otsu's thresholding to image (b). (d) Binary image after applying morphological conditions in order to avoid artefacts. Its contour is treated as the initial contour. (e) Enhanced edge indicator function. (f) Final segmentation. The final segmentation is indicated in red.

As the proposed pigmented lesion segmentation is part of a computer-aided classification system, probably to be used in real time, this restriction applies.

Table 3.8 Computational cost of computing the three colour gradients per image.

Computational cost	CIELAB	CIE94	CIEDE2000
(seconds/image)	0.69	1.81	3.1

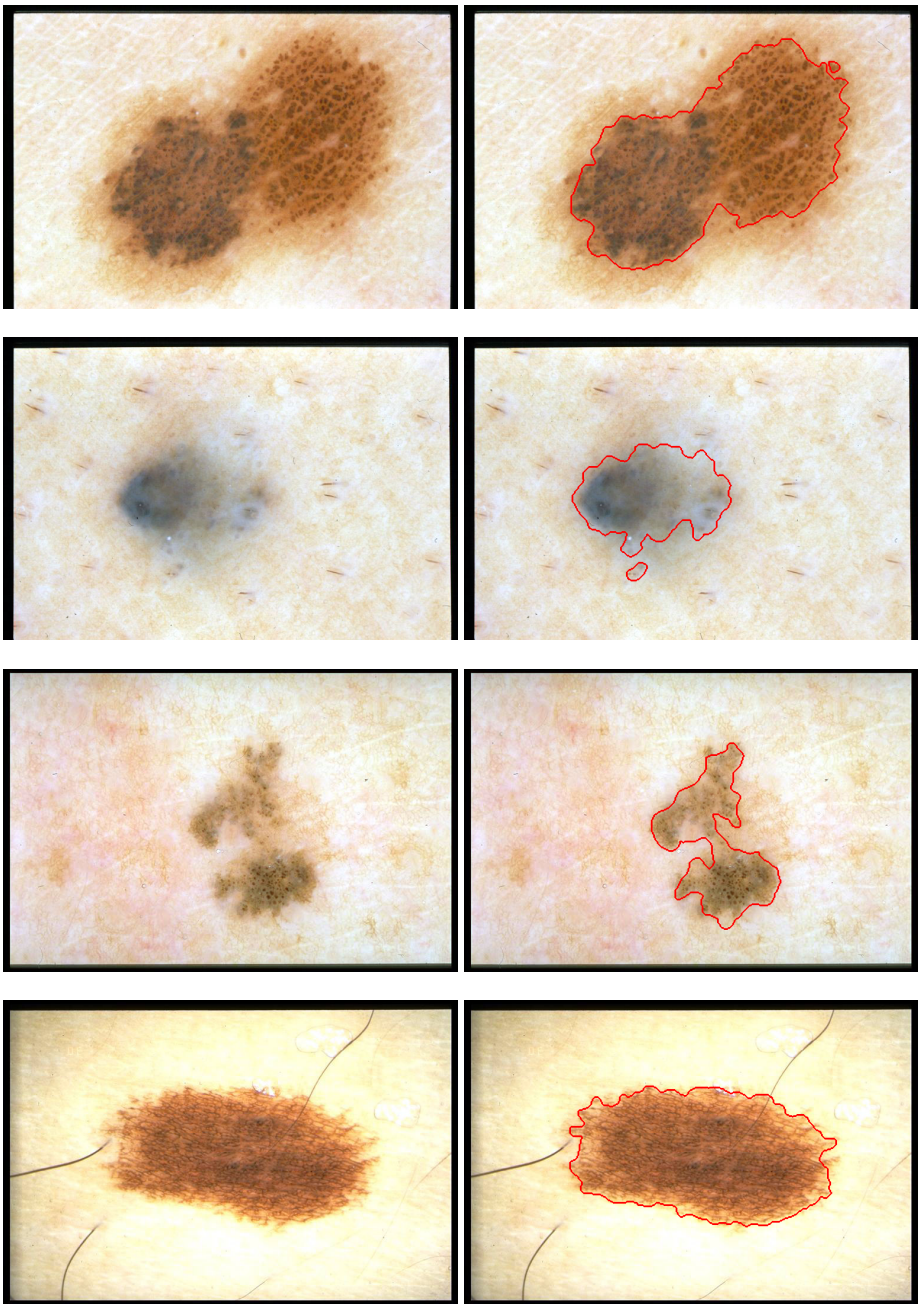


Figure 3.10 Examples of pigmented lesion segmented with the proposed algorithm.

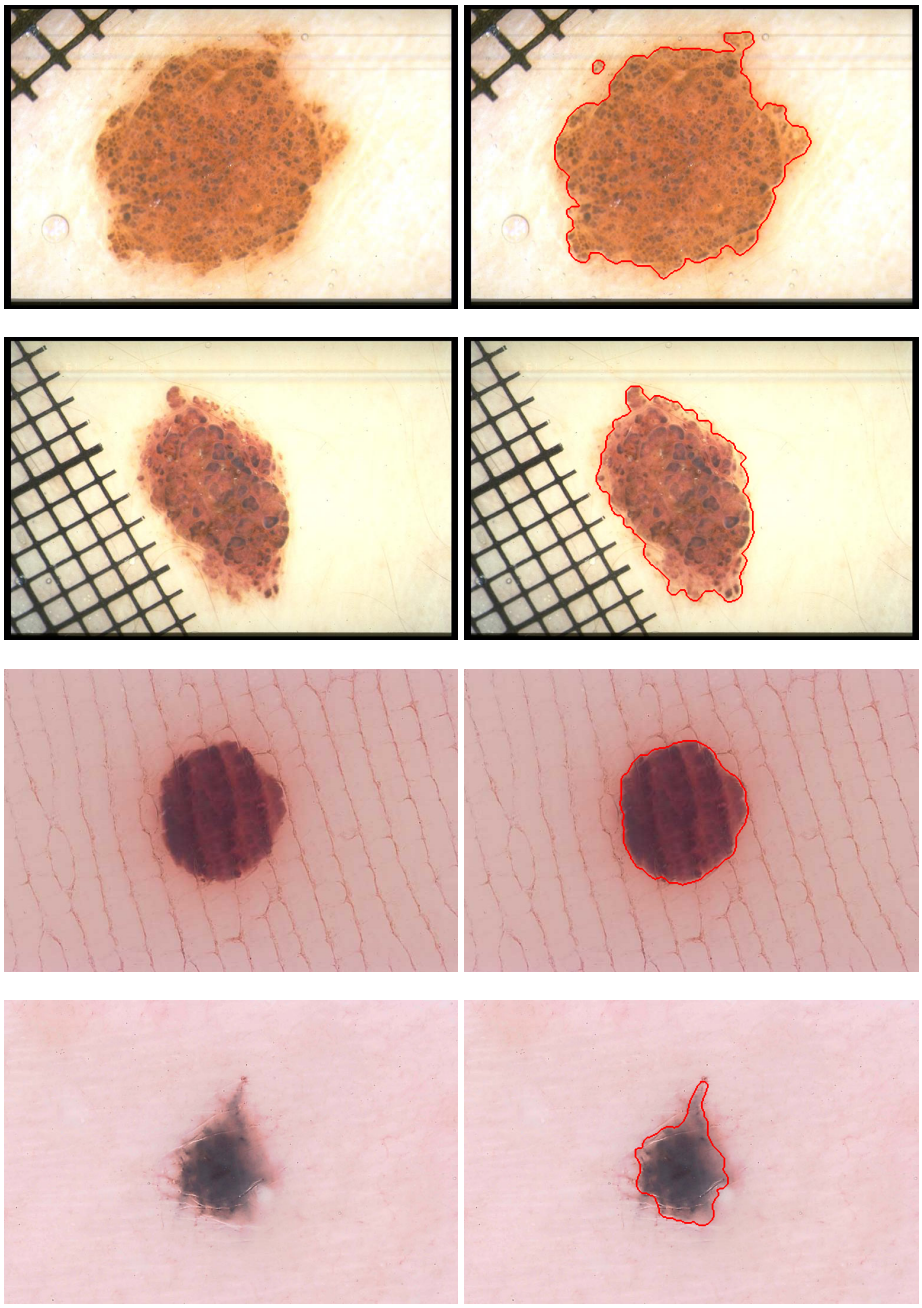


Figure 3.11 Examples of pigmented lesion segmented with the proposed algorithm .

CHAPTER 4

Related Publications

- Book chapters:

  Sáez A., Acha B., Serrano C. Pattern analysis in dermoscopic images. *In Computer Vision Techniques for the Diagnosis of Skin Cancer, published by Springer (Series in BioEngineering)* 2013. ISBN: 978-3-642-39607-6

4 Pattern analysis in dermoscopic images. Review

Never consider the study as an obligation but as an opportunity to penetrate the beautiful and wonderful world of learning.

Albert Einstein

Pattern Analysis is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma and it seeks to identify specific patterns, which may be global and local. In this chapter an extensive review of algorithmic methods found in the literature that automatically detect these patterns in dermoscopic images of pigmented lesions is presented.

4.1 Introduction.

The medical term melanoma refers to a malignant tumour developed from melanocytic cells. Melanoma generally appears de novo, and less frequently as the evolution of acquired benign melanocytic nevi. In the last decades, mainly due to sun exposure, the incidence of melanoma has dramatically increased, particularly in young white population. If diagnosed and treated early, the mean life expectancy of individuals suffering from melanoma can be increased by at least 25 years [Capdehourat et al., 2011].

A non-invasive technique to assist dermatologists in the diagnosis of melanoma is dermoscopy, which is an epiluminescence light microscopy, that magnifies lesions and enables examination down to the dermo-epidermal junction. There are four main diagnosis methods from dermoscopic images: ABCD rule, pattern analysis, Menzies method and 7-point checklist. These methods were evaluated during the 2000 Consensus Net Meeting on Dermoscopy (CNMD) [Argenziano et al., 2003] by experts from all over the world. A 2-step procedure was used to facilitate the diagnosis:

a) To decide whether the lesion is melanocytic or non-melanocytic.

b) To decide whether the melanocytic lesion is benign, suspect, or malignant.

Pattern Analysis, considered as the classic approach for diagnosis in dermoscopic images, was deemed superior to the other algorithms. The favourable results of pattern analysis were not unexpected, because this method probably reflects best the way the human brain is working when categorizing morphological images [Argenziano et al., 2003]. Pattern Analysis, set forth by Pehamberger and colleagues in 1987 [Pehamberger et al., 1987], was updated by the Consensus Net Meeting of 2000 [Argenziano et al., 2003]. This methodology defines the significant dermatoscopic patterns of pigmented skin lesions. Currently, it is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma [Rezze et al., 2006].

Pattern Analysis seeks to identify specific patterns, which may be global or local. The global features allow a quick preliminary categorization of a given pigmented skin lesion prior to more detailed assessment, and they are presented as arrangements of textured patterns covering most of the lesion. The local features represent individual or grouped characteristics that appear in the lesion. Some conclusions from this methodology were extracted in the consensus mentioned [Argenziano et al., 2003]; the global feature most predictive for the diagnosis of melanoma was the multicomponent pattern, whereas the globular, cobblestone, homogeneous, and starburst patterns were most predictive for the diagnosis of benign melanocytic lesions. Regarding to local features, atypical pigmented network, irregular streaks, and regression structures were the features that showed the highest association with melanoma, followed by irregular dots/globules, irregular blotches, and blue-whitish veil. Vascular structures were not found to be significantly associated with melanoma. On the contrary, typical pigmented network, regular dots/globules, regular streaks, and regular blotches were mostly associated with benign melanocytic lesions.

Although the 7-Point Checklist method corresponds to a different diagnostic technique than pattern analysis, it can be considered as a simplification of it, as it classifies 7 features related to local patterns [Argenziano et al., 1998]. Such simplified algorithm was designed to prevent non-experts from missing the detection of melanomas, even at the cost of decreased specificity.

Due to the proven benefits of applying digital imaging to dermatology [Stoecker and Moss, 1992, Argenziano et al., 2003], image processing research has directed a strong effort to develop Computer Aided Diagnosis (CAD) tools to assist physicians in their task of analysing pigmented lesions, especially because a dermatologist is not always the physician that analyses them in a first trial. In 2009 Maglogianis and Doukas [Maglogiannis and Doukas, 2009] presented an overview of CAD systems, describing how to extract features through digital image processing methods and techniques for skin lesion classification. The special issue -Advances in skin cancer image analysis [Emre Celebi et al., 2011]- summarized the progress that has taken place in this field, including works related to multispectral imaging system, enhancement of dermoscopy images, detection of lesion border and feature extraction. And the recent work from Korotkov and Garcia [Korotkov and Garcia, 2012] presents an extensive review of computerized analysis of pigmented skin lesions applied to microscopic (dermoscopic) and macroscopic (clinical) images.

In this chapter an exhaustive review of methods devoted to quantify features in pattern analysis is presented. In Section 4.2 a brief summary of the main techniques focused on the extraction of local patterns, including those that quantify the features related to 7-Point Checklist, is presented. In Section 4.3 the 7-Point Checklist method is detailed. When dealing with the detection and/or classification of global patterns, a few methods have been published. A description of these methods is introduced in Section 4.4.

4.2 Local Patterns

The presence of specific dermoscopic features in different regions of the same lesion contributes to make a diagnosis of melanocytic lesions and are called local patterns. They are dermoscopic structures such as pigment network, dots and globules, streaks, blue-whitish veil, regression structures, hypopigmentation and vascular structures, whose appearance description is presented below [Argenziano et al., 2000]. The predominant presence of some of these local patterns can determine some global patterns [Argenziano et al., 2003, Argenziano et al., 2000].

- Pigment network. Delicate, regular grid of brownish lines over a diffuse light-brown background.
- Dots/globules. Sharply circumscribed, usually round or oval, variously sized black, brown or grey structures.
- Streaks. Brownish-black linear structures of variable thickness
- Blue-whitish veil. Grey-blue to whitish-blue, diffuse pigmentation associated with pigment network alterations, dots/globules and/or streaks.
- Pigmentation. Dark-brown to grey-black, diffuse area that precludes recognition of more subtle dermoscopic features such as pigment network or vascular structures.
- Hypopigmentation. Diffuse area of decreased pigmentation within an otherwise ordinary pigmented lesion.
- Regression structures. White areas, blue areas and a combination of both. Virtually indistinguishable from the blue-whitish veil.
- Vascular structures.
- Other criteria, such as milia-like cysts, comedo-like openings, blotches, Lacunas network, etc.

Local patterns can be presented in the lesion with an irregular/regular or atypical/typical nature, implying malignancy or not. Fig.4.1 shows some examples of local patterns.

In the literature we can find numerous works that are focused on the automatic identification of local features. They are briefly explained in the subsections below.

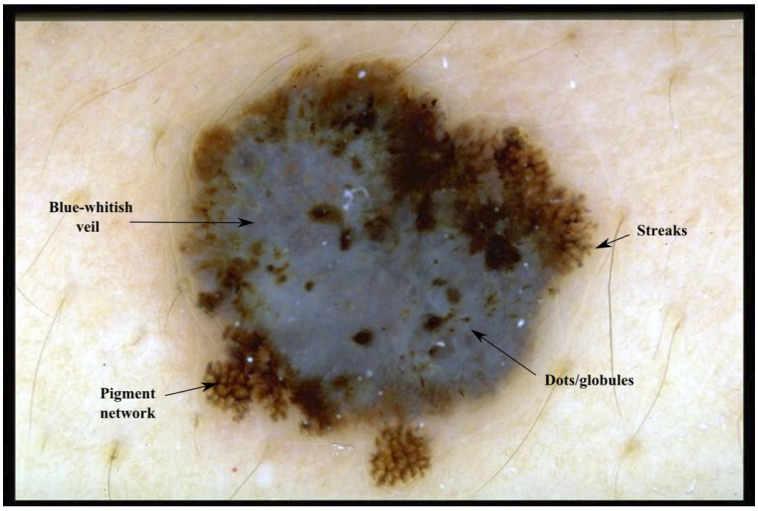


Figure 4.1 Example of local patterns..

4.2.1 Pigment network

The pigment network is the most studied local pattern. The reason is that it is the most common local pattern in melanocytic lesions, and the identification of melanocytic lesions is the first step in the procedure of pigmented skin lesion diagnosis. A pigment network can be typical, when the pattern is regularly meshed, narrowly spaced and its distribution is more or less regular, or atypical, characterized by a black, brown, or grey, irregular network, distributed irregularly throughout the lesion. An atypical network signals malignancy [Argenziano et al., 2000]. Fig. 4.2 shows the variability of its appearance.

In the last years several authors have focused on the automatic detection of this pattern.

Anantha et al. compared two statistical texture identification methods for detecting the pigment network [Anantha et al., 2004]. The first method was the neighbouring grey-level dependence matrix (NGLDM), and the second method used the lattice aperture waveform set (LAWS). They analysed images of 64x64 pixels. The authors concluded that both methods detect grossly any pigment network with reasonable accuracy, with slightly better results obtained by the latter. 155 dermatoscopic images were analysed, including 62 malignant melanomas and 93 benign lesions. The success classification percentage was around 78% and 65% for LAWS and NGLDM, respectively.

Fleming et al. [Fleming et al., 1998] and Grana et al. [Grana et al., 2006] presented an approach that addressed the problem of detecting the pigment network based on the Steger's work [Steger, 1998] where a method for linear structure identification is presented. Line points detection was satisfied considering the lines of the pigment network as ridges. As a consequence this set of points must satisfy at the same time two conditions: the first order derivative should be zero, while the second order derivative should

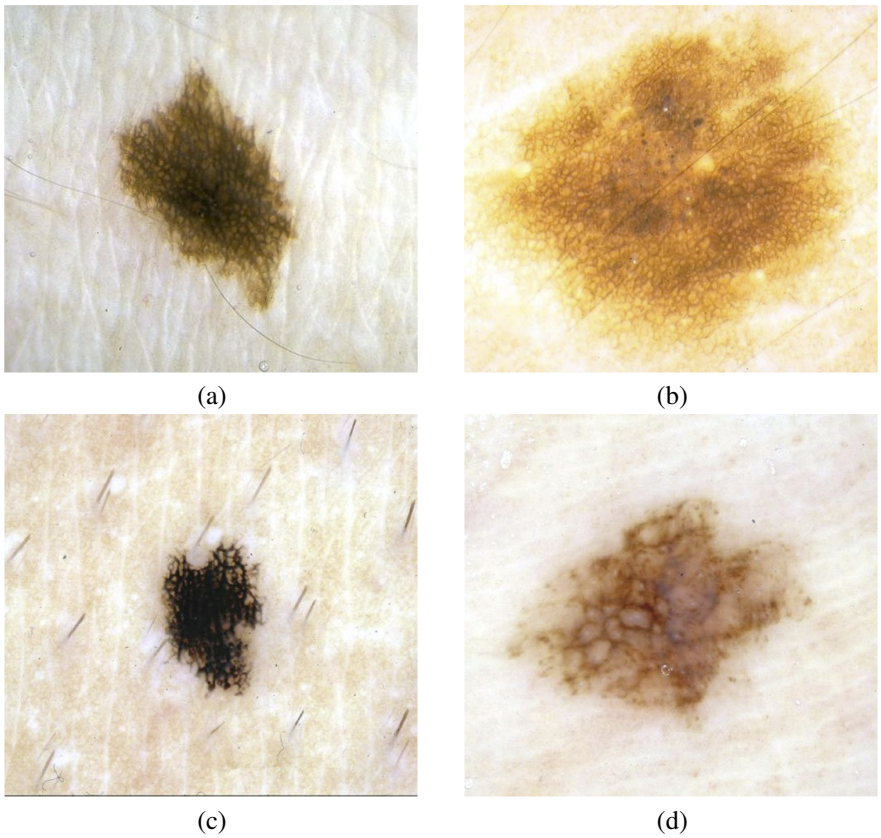


Figure 4.2 Example of lesions with pigment network. a) and b) present typical pigment network, whereas c) and d) atypical..

have a high module value. The final network is obtained linking these points. Fleming et al. [Fleming et al., 1998], following the procedure presented in [Steger, 1998], closed lines through an analysis of the second derivative, which gave information about the orientation of the line and the proximity between the points. However, in [Grana et al., 2006] a set of morphological masks that rotated in different directions in order to identify the terminations of the lines and, subsequently, line linking was used. After the network extraction, the image was divided into 8 sectors oriented along the principal axes, in order to provide some statistics on the network characteristics of the whole lesion and of every eight thereof counting the number of meshes, along with the number of unclosed terminations and the average line width. A set of 60 selected lesions was examined. Interestingly, the authors classified each lesion with regard to the distribution of the pigment network as no network pattern, partial network pattern if the lesion is partially covered with pigment network and complete network pattern. An overall 88.3% network detection performance,

without failed detections, was achieved.

Betta et al. proposed a method for the detection of atypical pigment network [Betta et al., 2006]. The method was based on their previous work [Betta et al., 2005], where the pigment network was detected but not classified as atypical/typical. The authors combined two different techniques: structural and spectral analysis. The structural analysis searched for primitive structures such as lines and/or points. To identify these local discontinuities the monochromatic image was compared with a median-filtered version of it and then a close-opening operation was applied. In the spectral analysis the Fourier transform of the monochromatic image was performed in order to determine the spatial period of the texture. In this way, local discontinuities, not clearly associated with the network, were disregarded. The result of this phase was a "regions with network" mask. This mask in conjunction with the structural mask provided a "network image", where the areas belonging to the lesion and constituting the pigment network were highlighted. Two indices related to the spatial and chromatic variability of these areas were presented to quantify the possible atypical nature of the network. 30 images were processed to assess the performance of this detection.

Leo et al. [Di Leo et al., 2008] extended the work proposed in [Betta et al., 2006] to detect atypical pigmented network. Once the pigment network was detected following [Betta et al., 2006], 13 colour and geometric features were extracted. C4.5 algorithm was used as classifier. 173 digital dermoscopy images (77 atypical pigment network, 53 typical pigment network and 43 absent pigment network) obtained from the Interactive Atlas of Dermoscopy [Argenziano et al., 2000] were used. 90 images were used for training and 83 images for testing. Sensitivity and specificity values greater than 85% were reached.

Shrestha et al. presented a study [Shrestha et al., 2010] whose purpose was to identify a method that could discriminate malignant melanoma with atypical pigment network (APN), from benign dysplastic nevi, which generally do not have an APN, using texture measurements alone. In this study, a gray-level co-occurrence matrix (GLCM) is constructed from the luminance plane. Five different GLCMs were constructed for each image using pixel distances (d-values) of 6, 12, 20, 30, and 40. Five classical statistical texture measures were calculated from each GLCM: energy, inertia, correlation, inverse difference, and entropy. Both the average and the range of each of these measures were computed, yielding 10 parameters related to texture. These parameters fed six different classifiers (BayesNet, ADTree, DecisionStump, J48, NBTree, and Ran- dom Forest) in order to determine whether an image presented pigment network or not. The method was tested with 106 dermoscopy images including 28 melanomas and 78 benign dysplastic nevi. The dataset is divided into APN areas and non-APN area. 10-fold validation is employed to validate the method. The "correlation average" provided the highest discrimination accuracy(95.4%). The best discrimination of melanomas is attained for a d-value of 20.

Sadeghi et al. [Sadeghi et al., 2010] proposed a method based on the detection of the "holes" of the network and follows the following steps: image enhancement, pigment network detection, feature extraction, and classification in three classes. First, a two-dimensional high-pass filter was applied to highlight texture features. Then, the lesion was segmented using Wighton et al.'s method [Wighton et al., 2009] which employed

supervised learning and the random walker algorithm. In the pigment network detection step, a Laplacian of Gaussian (LOG) filter was used to detect sharp changes of intensity. Then, the resulting binary image was converted into a graph using 8-connected neighbouring. Cyclic structures were found in this graph, and noise or undesired cycles were removed. Lines and holes of the pigment network were identified and 69 clinically inspired features were extracted: 20 structural features, including network thickness and its variation within the lesion, as well as size of the holes and its variation along the network; 2 geometric features to study the "uniformity" of the network; 37 chromatic features; and 10 textural features, using the five classical statistical texture measurements, also proposed in [Shrestha et al., 2010]. This allowed to classify the network into typical or atypical type. These 69 features were fed into a classifier based on a powerful boosting algorithm LogitBoost. A dataset consisting of 436 images (161 Absent, 154 Typical network, 121 Atypical network) was used. The authors computed results for both the 3-class (Absent, Typical or Atypical) and 2-class problems (Absent, Present). Ten-fold cross validation was used to generate all results. An accuracy of 82% discriminating between three classes and an accuracy of 93% discriminating between two classes were achieved. In [Sadeghi et al., 2011], the same authors, according to the density of the pigment network graph, classified a given image into Present or Absent. The method was evaluated with 500 images obtaining an accuracy of 94.3%.

Skrovseth et al. also proposed a pattern recognition technique with supervised learning to identify pigment network [Skrøvseth et al., 2010]. They selected a training set consisting of a large number of small images containing either a sample of network or a sample of other textures, either healthy skin or lesion regions. 20 different texture measures were analysed and the three that contributed maximally to separate the two classes with a linear classifier were selected. A new image is divided into overlapping subimages of the same size as the training images. A pixel is classified as network if at least one of the subimages it belongs to is classified as it.

Wighton et al. [Wighton et al., 2011] proposed the use of supervised learning and MAP estimation for automated skin lesion diagnosis. The authors applied this method to three task: segmentation, hair detection and identification of pigment network. The method was divided into three main steps. First, in a feature extraction stage, images were converted to CIE $L^*a^*b^*$ [Rangayyan et al., 2011], and each colour channel was filtered with a series of Gaussian and Laplacian of Gaussian filters at various scale ($\sigma = 1.25, 2.5, 5, 10, 20$), so that a total of 30 features were obtained for each pixel. Secondly, after feature extraction, Linear Discriminant Analysis (LDA) was used to reduce the dimensionality. Finally, the posterior probabilities $P(p|l_i)$ ($p = pixel$, $l_i = class$) in this subspace were modelled as multivariate Gaussian distributions. In the training phase, parameters for multivariate Gaussian distributions of each class were estimated. And in the labelling stage, individual pixels from previously unseen images were assigned a label using MAP estimation. A training dataset consisting in 20 images where pigment network was present across the entire lesion and 20 images absent of pigment network was employed. All the images belonged to the dermoscopy atlas presented in [Argenziano et al., 2000], where labels of "present" or "absent" of pigment network are supplied for each image. Pixels from the training images were assigned a label as "background", "absent" or "present". To label a new unseen image, features were computed as in the training phase and the dimensionality

of the feature space is reduced. To estimate the probability that a pixel p was labelled l_i, $P(l_i|p)$, the authors assigned the most probable label according to MAP estimation.

Barata et al. [Barata et al., 2012] presented a work based on the use of directional filters. The first step was to convert the colour image into a grey scale one to remove two types of artefacts: hair and reflections caused by the dermatological gel. An inpainting technique was applied. In the second step, regions with pigment network were detected using two of its distinctive properties: intensity and geometry or spatial organization. A bank of directional filters was applied to perform an enhancement of the network. The spatial organization was implemented by connectivity among pixels. The result was a binary net-mask. The final step aimed to assign a binary label to each image: with or without pigment network. To accomplish this objective, features which characterize the topology of the detected regions in a given image were extracted and used to train a classifier using a boosting algorithm. The algorithm was tested on a dataset of 200 dermoscopic images (88 with pigment network and 112 without) achieving a sensitivity of 91.1%, a specificity of 82.1% and an accuracy of 86.2% in the classification with or without pigment network.

In Table 4.1 the classification results of the works reported in this section are summarized.

Table 4.1 Results of **pigment network** detection..

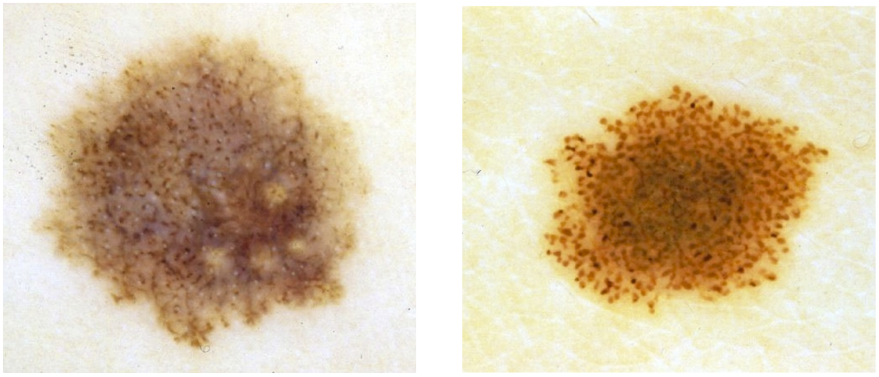
Algorithm	Classification	Accuracy	No. images
[Anantha et al., 2004]	Absent/Present	78%	155
[Grana et al., 2006]	No/Partial/Complete	88.3%	60
[Betta et al., 2006]	Atypical/Typical	-	30
[Di Leo et al., 2008]	Atypical/Typical	85%	173
[Shrestha et al., 2010]	Melanoma/No	95.4%	106
[Sadeghi et al., 2010]	Absent/Present	93%	436
[Sadeghi et al., 2010]	Absent/Atypical/Typical	82%	436
[Skrøvseth et al., 2010]	Absent/Present (per-pixel)	-	-
[Wighton et al., 2011]	Absent/Present (per-pixel)	-	734
[Barata et al., 2012]	Absent/Present	86.2%	200

4.2.2 Dots and Globules

Dots and globules are round or oval, variously sized black, brown or grey structures. It is another dermoscopic structure which is difficult to discriminate from pigment network [Sadeghi et al., 2010]. This could be the reason why there are so few works in the literature focused on its identification. Some examples of lesions with this structures are shown in Fig. 4.3.

Yoshino et al. [Yoshino et al., 2004] presented an algorithm that used morphological closing operation to detect dots. The closing operation used a linear structural element. Afterwards, a thresholding is applied to detect dots.

Based on the classification described in [Ojala et al., 2002], [Skrøvseth et al., 2010]



(a) (b)

Figure 4.3 Example of lesions with dots/globules..

computed a score for each pixel in a gray scale image. Given P surrounding pixels with values g_k, $k = 1,...,P$ at a radius R of a central pixel with gray value g_c, the score of the central pixel is calculated as $S_c = \Sigma_{k=1}^{P}(g_c - g_k)$. The authors argue that this score will be large for a dark spot, and therefore, a simple thresholding would give the position of the dot.

Table 4.2 shows a summary of the works focused on the globules detection.

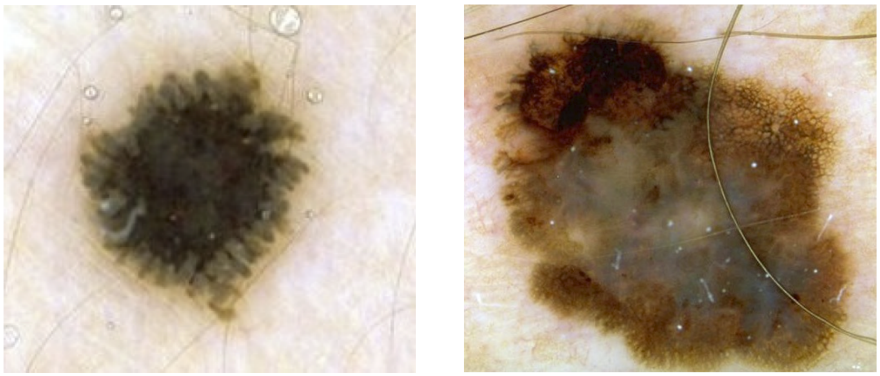
Table 4.2 Results of **dots/globules** detection..

Algorithm	Classification	Accuracy	No. images
[Yoshino et al., 2004]	Absent/Present	-	-
[Skrøvseth et al., 2010]	Absent/Present	-	-

4.2.3 Streaks

Streaks are brownish-black linear structures of variable thickness that are found in benign and malignant lesions. They are typically placed at the periphery of a lesion and are not necessarily connected to the lines of the pigment network. Streaks can be irregular, when they are unevenly distributed (malignant melanoma), or regular (symmetrical radial arrangement over the entire lesion) [Braun et al., 2005]. An example of regular and irregular streaks can be found in Fig. 4.4.

Beta et al. identified streaks as finger-like irregularities with uniform brown colour at the lesion contour [Betta et al., 2005]. Therefore, they detected the simultaneous occurrence of two different structures: finger-like track of the lesion contour, and brown pigmentation in the same restricted region. For the first purpose, the colour image was converted to an 8-bit gray-level image, and then three different binary images were obtained



(a) (b)

Figure 4.4 Example of lesions with a) regular streaks b) irregular streaks..

by applying three different thresholds. The contours of these binary images were extracted by a blob-finding algorithm. The best of the three extracted contours was selected manually and divided into 16 parts. For each part, an irregularity ratio was evaluated. This parameter represented the ratio between number of pixels of the detected contour in this part and the number of pixels in the line connecting the extreme contour points of this part. The contour in that region was assumed as irregular if the ratio was greater than a threshold. On the other hand, brown pigmentation of those 16 subimages is analysed by thresholding the hue component. Finally, the occurrence of streaks was assumed only if both an irregular contour and a brown pigmentation were found in the same sub-image. The authors presented experimental results for 10 images achieving a 90% of success rate. A further evaluation was presented by Fabbrocinia et al. [Fabbrocinia et al., 2010]. They used 23 and 30 images for training and test set, respectively. The two thresholds mentioned above were determined by a Receiver Operating Characteristic curve (ROC curve) on the training image set. A sensitivity and a specificity of 86% and 88%, respectively, were achieved.

A machine-learning approach to detect streaks which captures the quaternion tubularness in the colour dermoscopic images was proposed in [Mirzaalian et al., 2012]. First, tubularness filters [Frangi et al., 1998] to enhance streak structures in dermoscopic images were used. Given the estimated tubularness and direction of the streaks, a vector field in order to quantify radial streaming pattern of the streaks was defined. Specifically, they computed the amount of flux of the field passing through iso-distance contours of the lesion, where each contour was the loci of the pixels which have equal distance from the outer lesion contour. So, an appearance descriptor based on the mean and variance of the flux through the different concentric bands of the lesion is constructed. The final step is to learn how the extracted descriptors can best distinguish the three different classes: the absence, presence of regular, or presence of irregular streaks in the dermoscopic images. This task is performed with a SVM classifier with a database 99 dermoscopic images.

In [Sadeghi et al., 2012b], the authors followed four steps to locate streaks: preprocess-

ing, blob detection, feature selection and two-class classification (absent-present). In the preprocessing step, lesions were segmented, reoriented so that the major axis was parallel to the x-axis and resized so that its major axis occupied 500 pixels. Lightness component (L*) from L*a*b* colour representation was used for the rest of the analysis. Streaks can be modelled as linear structures with a Gaussian cross-section profile near the border. Therefore, in the blob detection, four Laplacian of Gaussian (LOG) filters with different sizes, hsize = 3, 5, 7, 9, were employed to detect these linear structures. Candidate to streaks were extracted in this step. Once they were detected, their orientations were estimated using the Averaged Squared Gradient Flow (ASGF) algorithm [Kass and Witkin, 1987]. Finally, 25 features were extracted from the candidate linear streak structures and from the lesion: one set of 12 features was based on properties of the detected candidate streak lines and another feature set contained the 13 common colour and texture features of the entire lesion. These 25 features were fed to a SimpleLogistic classifier, that classifies a lesion into absence and presence of streaks. The method was tested with a database of 300 dermoscopic images (105 Absent and 195 Present) achieving an accuracy detection of 0.815 using 10-fold cross validation.

Sadeghi et al. [Sadeghi et al., 2013] recently presented an extension of their previous work [Sadeghi et al., 2012b]. In this version, they proposed an algorithm that classifies a lesion into absence of streaks, regular streaks, and irregular streaks. The work aimed to identify valid streak lines from the set of candidate streak lines obtained in [Sadeghi et al., 2012b] in order to reduce false positive streaks such as hairs and skin lines. The method also extends the analysis to identify the orientation and spatial arrangement of streak lines. These novel geometric features are used to identify not only the presence of streak lines, but whether or not they are Irregular or Regular; important for melanoma diagnosis. Therefore, a total of 31 features are fed into a classifier, achieving an accuracy of 76.1% when classifying 945 images into the three classes.

Table 4.3 summarizes the classification results of the works reported in this section.

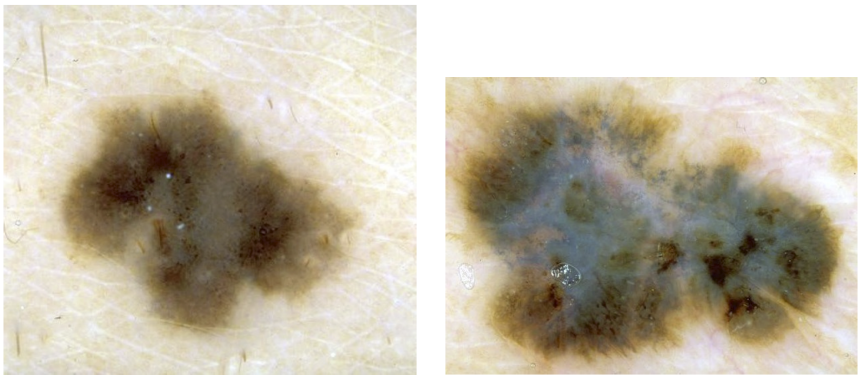
Table 4.3 Results of **streaks** detection..

Algorithm	Classification	Accuracy	No. images
[Betta et al., 2005]	Absent/Present	90%	10
[Fabbrocinia et al., 2010]	Absent/Present	86%	30
[Mirzaalian et al., 2012]	Absent/Regular/Irregular	91%	99
[Sadeghi et al., 2013]	Absent/Present	78.3%	945
[Sadeghi et al., 2013]	Absent/Regular/Irregular	76.1%	945

4.2.4 Blue-whitish veil

Blue-whitish veil is characterized by a grey-blue to whitish-blue diffuse pigmentation. Some examples of lesions that present this pattern are shown in Fig. 4.5.

Celebi et al. [Celebi et al., 2008a] proposed a machine learning approach to detect blue-white veil in dermoscopy images based on a previous work [Celebi et al., 2006]. Fifteen



(a) (b)

Figure 4.5 Example of lesions with blue-whitish veil..

colour features and three texture features were extracted. The colour features involved absolute colour features and relative colour features when compared to the average colour of the background skin. The texture features were based on the Gray Level Co-occurrence Matrix (GLCM). The classifier used was C4.5 algorithm. Only 2 out of the 18 features were finally selected for the classification model, both belonging to colour features. The classification results for manually selected test pixels yield a sensitivity of 84.33% and a specificity of 96.19%. In a second experiment, the authors aimed to discriminate between melanoma and benign lesions based on the area of the blue-white veil detected. They extracted three numeric values from the detected blue-white region: area, circularity and ellipticity. A new classification model based on these features was generated using C4.5 algorithm and 10-fold-cross validation. A sensitivity of 69.35% and a specificity of 89.97% for the entire image set (545 images) were obtained.

Leo et al. [Di Leo et al., 2009] focused on the detection of two different patterns, blue-whitish veil and regression structures. Firstly, the lesion was subdivided into regions. The colour image is converted via Principal Component Analysis (PCA) and a two dimensional (2-D) histogram is computed with the two first principal components. The most significant peaks in the 2-D histogram were found as representative colour in the input image. All the pixels in the lesion were assigned to one of the main peaks via clustering, so that a lesion map was created. Regions in the lesion map were subsequently classified as present or absent of blue whitish veil and regression. To this aim geometric and colour features were extracted and a Logistic Model Tree (LMT) was proposed as classifier. 210 digital dermoscopic images obtained from the Interactive Atlas of Dermoscopy [Argenziano et al., 2000] were used. 70 and 50 cases corresponding to the presence of Blue Veil and Regression area respectively were used as training set. 65 cases of Blue Veil and 40 cases of Regression structures were utilized for the test set. A sensitivity of 0.87 and a specificity of 0.85 were obtained for the detection of blue veil and a sensitivity and a specificity both equal to 0.85 for regression structures.

In a recent work, Arroyo et al. [Arroyo et al., 2011] also proposed supervised machine

learning techniques to detect blue-white veil. To this aim, colour features were extracted from each individual pixel and the classifier used was C4.5 algorithm, that generated the decision tree. The authors used a database consisting of 887 images. 120 images were selected to obtain the training data of the machine learning algorithms, 60 corresponding to melanoma with the blue-white veil pattern and 60 corresponding to other cases. Colour features were extracted from the pixels. After generating the pixel rules, these are applied to all the images to detect candidate areas. A new classification model was generate to obtain the rules that must satisfy the whole image to be a melanoma with blue-white veil. For this purpose 12 features were extracted from the candidate area such as area, solidity or ellipticity. The method achieved a sensitivity of 80.50% and a specificity of 90.93% when the 887 images were classified.

Table 4.4 shows the classification results of the mentioned works.

Table 4.4 Results of **blue-whitish veil** detection.

Algorithm	Classification	Sensitivity	Specificity	No. images
[Celebi et al., 2008a]	Absent/Present	84.33%	96.19%	100
[Celebi et al., 2008a]	Melanoma/Benign	69.35%	89.97%	545
[Di Leo et al., 2009]	Absent/Present	87%	85%	135
[Arroyo et al., 2011]	Melanoma/Benign	80.50%	90.93%	887

4.2.5 Blotches

Blotches are dark structureless areas within pigmented lesions [Argenziano et al., 2003]. Blotches that are located asymmetrically within a lesion are indicative of malignant melanoma [Khan et al., 2009].

Stoecker et al. [Stoecker et al., 2005] studied the effectiveness of the absolute and relative colour blotch features for melanoma/benign lesion discrimination over a dermoscopy image set containing 165 melanomas and 347 benign lesions using a neural network approach. The authors proposed two approaches to detect the blotchy areas. The first method used thresholds placed upon the values of the red, green, and blue (RGB) components of the pixels within the lesions. The second method used relative colour thresholds, subtracting the observed pixel value within the lesion from the background skin colour before applying relative thresholds. Then, several blotch indices were computed, including the scaled distance between the largest blotch centroid and the lesion centroid, ratio of total blotch areas to lesion area, ratio of largest blotch area to lesion area, total number of blotches, size of largest blotch, and irregularity of largest blotch. It was determined that relative colour were more effective than absolute colour giving a diagnostic accuracy of about 77%.

In [Khan et al., 2009] new and existing blotch features for melanoma discrimination are investigated. Four experiments were performed to achieve this aim. Blotch candidates are first extracted using absolute and relative colour thresholds to construct blotch masks proposed in [Stoecker et al., 2005]. Then, fuzzy logic techniques for extracting blotches

based on blotch size were studied, where a fixed minimum blotch size was fuzzified to detect an area candidate as a blotch only if its size exhibits a certain degree of association with a fuzzy set representative of blotch size. To compute the second fuzzy set, the relative colour values at each pixel position inside of the blotchy areas were extracted from melanoma lesions belonging to a training set of images. This fuzzy set provided the basis for differentiating between melanoma and benign skin lesions. A third fuzzy set was constructed similarly, but using separate relative colour histograms for the red, green and blue colour planes. These sets were also used for melanoma discrimination. Finally, a new set of four asymmetry features were computed. The lesion border mask was divided into four quadrants and a set of asymmetry features was computed. The work concluded that features computed from blotches using the fuzzy logic techniques based on three plane relative colour and blotch size yielded the highest diagnostic accuracy of 81.2%. 424 dermoscopy images (134 melanoma and 290 benign images) from three different sources were used.

Madasu and Lovell [Madasu and Lovell, 2009] proposed an extension of Fuzzy Co-Clustering Algorithm for Images (FCCI) technique [Hanmandlu et al., 2008] for detecting blotches. Madasu and Lovell extended FCCI technique to include texture features as additional clustering parameters. Texture features were computed using the normalized entropy function. A set of 50 images were used for testing the proposed algorithm. The authors claimed that the blotches are accurately located independently of their shape, size or location within the image.

A summary of classification results of the three works presented is shown in Table 4.5.

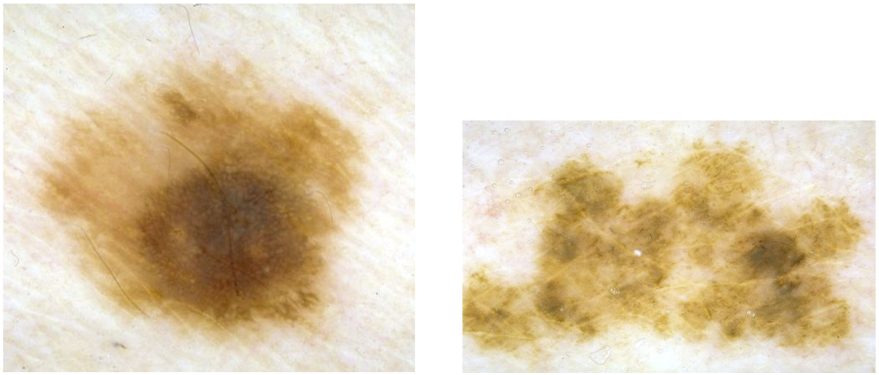
Table 4.5 Results of **blotches** detection..

Algorithm	Classification	Accuracy	No. images
[Stoecker et al., 2005]	Melanoma/Benign	77%	512
[Khan et al., 2009]	Melanoma/Benign	81.2%.	424
[Madasu and Lovell, 2009]	Absent/Present	-	50

4.2.6 Hypopigmentation

Hypopigmentation represents a diffuse area of decreased pigmentation within an otherwise ordinary pigmented lesion. White areas in a melanoma tend to have a central position. White areas in a nevus are located in the lesion periphery.

Dalal et al. [Dalal et al., 2011] proposed a method to discriminate melanomas from benign nevi by automatically detecting white areas and measuring features of these white areas. In order to identify white and hypopigmented areas, thresholds were determined for each colour plane based on colour histogram analysis over a training set of images. The lesion was segmented in concentric deciles. Overlays of white areas on the lesion deciles were determined. Nine indices were calculated to characterize the automatically detected white areas in a lesion. These indices included lesion decile ratios, normalized number of white areas, absolute and relative size of largest white area, relative size of all



(a) (b)

Figure 4.6 a) Nevu with hypopigmentation. b) Melanoma with hypopigmentation..

white areas, and white area eccentricity, dispersion, and irregularity. A neural network was selected as classifier. 244 benign and malignant dermoscopy images with white areas were selected. The methods used a randomly selected training set of 75 lesions and a test set of 169 lesions.

Table 4.6 Results of **hypopigmentation** detection..

Algorithm	Classification	AUC	No. images
[Dalal et al., 2011]	Melanoma/Benign	95.2%	244

4.2.7 Regression structures

A regression structure is a white scarlike depigmentation irregularly distributed within the lesion.

As mentioned above, in [Di Leo et al., 2009] regression structures and blue-whitish veil were detected. The method is reported in Section 4.2.4. Table 4.7 shows the classification result of this work.

Table 4.7 Results of **regression structures** detection..

Algorithm	Classification	Sensitivity	Specificity	No. images
[Di Leo et al., 2009]	Absent/Present	85%	85%	90

4.2.8 Vascular pattern

A vascular pattern, and more specifically, with atypical nature presents linear-irregular or dotted vessels not clearly combined with regression structures and associated with pigment network alterations, dots/globules and/or streaks.

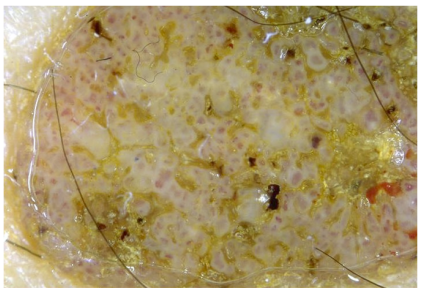


Figure 4.7 Example of lesion with dotted vessels..

In [Betta et al., 2006], reported in Section 4.2.1, a method for the detection of atypical vascular pattern is also proposed. Due to the difficulty to obtain a relevant number of ELM images with the occurrence of this criterion, the training set was constituted by N pixels selected as vascular patterns in a set of images containing occurrences of this criterion. The Hue, Saturation and Luminance components was evaluated and the frequency histograms corresponding to the three colour planes were determined. The pixel classification depended on the value of its HSL components. However, the authors warned that in some cases the algorithm gave rise to wrong detection, evidencing a low specificity.

4.3 7-Point Checklist method

Although the 7-point checklist method is a different diagnosis algorithm, it is considered a simplification of the classic pattern analysis due to the low number of local features to identify. This algorithm is applied once the lesion is diagnosed as melanocytic. It was developed by Argenziano et al. [Argenziano et al., 1998]. The 7-Point Checklist is a score system. This method uses seven specific criteria for melanoma. It includes three major criteria:

- Atypical pigment network
- Blue-whitish veil
- Atypical vascular pattern

2 points are attributed to each of the former, and four minor criteria:

- Irregular streaks

- Irregular pigmentation
- Irregular dots/globules
- Regression structures

1 point is attributed to each of them. A score of 3 or greater is associated with a high likelihood of melanoma diagnosis at pathology evaluation.

Some works focused on the detection of one or several specific criteria can be found in the literature [Betta et al., 2006, Di Leo et al., 2008, Betta et al., 2005, Fabbrocinia et al., 2010, Di Leo et al., 2009]. These works have been explained in Section 4.2.

Di Leo et al. in [Di Leo et al., 2010a] and [Di Leo et al., 2010b], joined some methods also mentioned in Section 4.2 in order to present an automatic implementation of the 7-Point Checklist method. In [Di Leo et al., 2010a] the authors focused on the detection of five criteria. It is in [Di Leo et al., 2010b] where the seven criteria of the method were addressed: Atypical pigment network and Irregular dots/globule were detected by the methods proposed in [Di Leo et al., 2008] and [Betta et al., 2006]; Blue-whitish veil, Regression structures and Irregular pigmentation detection followed the steps proposed in [Di Leo et al., 2009]; atypical vascular pattern was detected with [Betta et al., 2006]; and Irregular streaks with the method presented in [Betta et al., 2005]. 300 images were used for the evaluation. For each dermoscopic criterion a training and testing set were selected from the database in order to train a classifier and/or carry out a statistical analysis. The system distinguished between Melanoma and Benign lesions. The performance of the automatic system was estimated through a comparison with the application of the 7-Point Checklist diagnostic method by expert dermatologists to 287 images of the database. The global sensitivity and specificity values of the software diagnostic tool were 0.83 and 0.76, respectively.

Other authors [Iyatomi et al., 2007], [Capdehourat et al., 2011] focused their works on machine learning approaches, where the feature extraction step was inspired on the 7 point checklist criteria.

4.4 Global pattern

Global features allow a quick preliminary categorization of a given pigmented skin lesion prior to more detailed assessment, and they are presented as arrangements of textured patterns covering most of the lesion.

The main global patterns are [Argenziano et al., 2000]:

- Reticular pattern. The most common global feature in melanocytic lesions is characterized by a pigment network covering most parts of a given lesion. Basically, the pigment network appears as a grid of thin brown lines over a diffuse light brown background.
- Globular pattern. It is characterized by the presence of numerous, variously sized, round to oval structures with various shades of brown and grey-black coloration.

- Cobblestone pattern. It is quite similar to the globular one but is composed of closely aggregated, larger somewhat angulated globules resembling a cobblestone.
- Homogeneous pattern. It appears as a diffuse, brown, grey-blue to grey-black or reddish-black pigmentation in the absence of pigmented network or other distinctive local features.
- Starbust pattern. It is characterized by the presence of pigmented streaks in a radial arrangement at the edge of a given pigmented skin lesion.
- Parallel pattern. It is found exclusively in melanocitic lesions on skin of palms and soles due to particular anatomic structures inherent to this location.
- Multicomponent pattern. It is a combination of three or more distinctive dermoscopic structures within a given lesion.
- Lacunar pattern. It is characterized by various to numerous, smooth-bordered, round to oval, variously sized structures called red lacunas, whose morphologic hallmark is their reddish, blue-purplish or black coloration.
- Unspecific pattern. In some instances, a pigmented lesion cannot be categorized into one of the global patterns listed above, because its overall morphologic aspect does not fit at all into these artificial, albeit rather distinctive categories. For this type of lesion the term unspecific pattern is used.

In [Tanaka et al., 2004] and [Tanaka et al., 2008] an analysis of texture classified a pattern into three categories: homogeneous, globular and reticular. The lesion area was divided into small regions. 110 texture features of each sub-image were calculated. These features were based on intensity histogram information, differential statistical features, Fourier power spectrum, run-length matrix, coocurrence matrix and connected components. 35 features were selected by discriminant analysis. As a result, the patterns were classified into three categories with a 94% of classification success rate.

The work [Iyatomi et al., 2008] is only focused on the detection of parallel pattern. 428 image features were extracted, which included colour-related features, symmetry features, border-related features and texture features. Then, using principal component analysis (PCA), these features were transformed into 198 orthogonal principal components (PCs) without information loss. The first 10 PCs more discriminative were selected. Four linear classifiers were used for parallel ridge, parallel furrow and fibrillar pattern detection. In addition, acral volar melanoma was also classified. The achieved results are shown in Table 4.8.

In [Gola Isasi et al., 2011], the authors presented a method which, in conjuction with the ABCD rule, tried to detect three global patterns in order to increase diagnostic accuracy of pigmented lesions. To this aim they developed three different algorithms. In the two first ones, based on edge detection and mathematical morphology, they detected globular and reticular patterns. In the third one, they performed colour analysis in the RGB colour space with the aim of detecting the blue veil pattern. The algorithms were tested with a database consisting of 20 images per global pattern. The proposed algorithms produced an average accuracy above 85%.

Table 4.8 Classification results in [Iyatomi et al., 2008]. SE=Sensitivity, SP=specificity, AUC=Area under the Receiver operating characteristic (ROC) curve.

Classifier	SE(%)	SP(%)	AUC
Melanoma	100	95.9	0.993
Parallel ridge pattern	93.1	97.7	0.985
Parallel furrow pattern	90.4	85.9	0.931
Fibrillar pattern	88	77.9	0.89

Table 4.9 Classification results in [Abbas et al., 2013]. SE=Sensitivity, SP=specificity, AUC=Area under the Receiver operating characteristic (ROC) curve.

Pattern	SE(%)	SP(%)	AUC
Reticular	87.11	97.96	0.981
Globular	86.25	97.21	0.997
Cobblestone	87.76	93.23	0.990
Homogeneous	90.47	95.10	0.996
Parallel	85.25	89.50	0.989
Starburst	89.62	90.14	0.966
Multicomponent	98.50	93.11	0.989

In their work, Abbas et al. [Abbas et al., 2013] extracted colour and texture features from a dermoscopic image in order to classify it into its global pattern. Colour related images were extracted from the CIECAM02 representation of the colour image. Texture was analysed by means of the steerable pyramids transform (SPT). Both groups of features fed an AdaBoost MC classifier, which classified pigmented lesions into seven different groups of global patterns: (a) Reticular pattern or pigmented network, (b) Globular pattern, (c) Cobblestone pattern, (d) Homogeneous pattern, (e) Parallel pattern, (f) Starburst pattern, (g) Multicomponent pattern. In Table 4.9 results are summarized.

Serrano and Acha [Serrano and Acha, 2009] were pioneers in the classification of global patterns following a model-based technique. They proposed a method to automatically classify five types of global patterns (reticular, globular, cobblestone, homogeneous and parallel), in which a Markov random field based texture modelling was performed. This method is detailed in Section 5.2 and Section 5.4 of Chapter 5.

In [Sadeghi et al., 2012a] the authors also detected and classified five classes of global lesion patterns (reticular, globular, cobblestone, homogeneous, and parallel ones). To this purpose, an approach based on texton classification [Varma and Zisserman, 2003] was followed, where texture features were modelled by the joint probability distribution of filter responses. The texton-based classification was performed in the L*a*b* colour space and in the gray-scale image using different filter banks. The procedure was divided into two steps:

Learning stage.

a) A set of 81 x 81 pixel images representing the five patterns were assembled.

b) Training images were convolved with a filter bank to generate filter responses.

c) Exemplar filter responses were chosen as textons (via K-Means clustering) and were used to label each filter response in the training images.

d) The histogram of texton frequencies was used to form models corresponding to the training images.

Classification stage.

a) The same procedure as in the training stage was followed to build the histogram corresponding to the unseen image.

b) This histogram was then compared with the models of the texton dictionary.

c) A nearest neighbour classifier was used and the Chi-square statistic was employed to measure distances.

The proposed set of filters was a filter bank composed by $18+18+2$ filters to detect average intensity, edges, spots, wave, meshes and ripples of dermatoscopic structures. The correct classification rate attained was 86.8%.

4.5 Discussion

Pattern analysis is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma [Rezze et al., 2006]. In fact, it was deemed superior to the other algorithms (i.e., ABCD Dermoscopy Rule, Menzies score, 7-Point Checklist) for diagnostic efficiency by experts from all over the world in the 2000 Consensus Net Meeting on Dermoscopy (CNMD) [Argenziano et al., 2003].

Pattern analysis aims to detect local or global patterns in a pigmented lesion to determine if it is melanocytic and, in such a case, its malignancy [Argenziano et al., 2003]:

a) Detection of pigment network, aggregate globules, streaks, homogeneous blue or parallel pattern are signs of melanocytic lesions.

b) Atypical pigment network, dots or streaks irregularly distributed, blue-white veil or regression may be signs of melanoma.

The presence of specific dermoscopic features in different regions of the same lesion contributes to make a diagnosis of melanocytic lesions and these features are called local patterns. The predominant presence of some of these local patterns determines some global patterns.

In the literature, many works devoted to detect local patterns have been published. Most of the papers related to local patterns address the problem of the detection of the pigmented network, which is the most common local pattern appearing in melanocytic lesions. However, there are only a few works that addressed the global patterns detection. The next chapter will be devoted to this issue.

CHAPTER 5

Related Publications

- International journals:

  Sáez A., Serrano C., Acha B. Model-based Classification Methods of Global patterns in dermoscopic images. *Submitted to IEEE Transactions on Medical Imaging.*

5 Model-based classification methods of global patterns

Divide each difficulty into as many parts as is feasible and necessary to resolve it.

RENE DESCARTES, 1637

Different methods of classification of global patterns in dermoscopic images are proposed in this chapter. These methods are based on modelling, applying this concept in two senses: first a dermoscopic image is modelled by a finite symmetric conditional Markov model (FSCM) applied to $L^*a^*b^*$ colour space and the estimated parameters of this model are treated as features. In turn, the distribution of these features are supposed that follow different models along a lesion: a Gaussian model, a Gaussian mixture model and a bag-of-features histogram model. For each case, the classification is carried out by an image retrieval approach with different distance metrics. The objective is to classify a whole pigmented lesion into three possible patterns: globular, homogeneous and reticular. An extensive evaluation of the performance of each method has been carried out on an image database extracted from a public Atlas of Dermoscopy. The best classification success rate is achieved by the Gaussian mixture model-based method with a 78.44% success rate in average.

5.1 Introduction.

As it has been mentioned in Chapter 4, pattern analysis is the method most commonly used for providing diagnostic accuracy for cutaneous melanoma and it seeks to identify specific patterns, which may be global and local. Numerous works have focused on the extraction of local patterns (Section 4.2 in Chapter 4), however, when dealing with the detection and/or classification of global patterns, a few methods have been published in the literature (Section 4.4 in Chapter 4). Tanaka et al. [Tanaka et al., 2008] presented an

extraction of 110 texture features to classify a pattern into three categories: homogeneous, globular and reticular. Gola et al. [Gola Isasi et al., 2011] presented a method based on edge detection, mathematical morphology and colour analysis to detect three global patterns (reticular, globular and homogeneous), but based on the predominant local pattern identification: globules, pigment network and blue pigmentation. Abbas et al. [Abbas et al., 2013] extracted colour features from the CIECAM02 representation and texture features from steerable pyramids transform (SPT) from the dermoscopic image in order to classify it into the seven global patterns. Serrano and Acha [Serrano and Acha, 2009] were pioneers in the classification of global patterns following a model-based technique. They proposed a method to automatically classify five types of global patterns (reticular, globular, cobblestone, homogeneous and parallel), in which a Markov random field based texture modelling was performed. Lately, Sadeghi et al. [Sadeghi et al., 2012a] modelled the texture with the joint probability distribution of filter responses to detect five patterns. However, these works classify patches extracted from a lesion instead of a whole lesion. To the best of our knowledge only Mendoza et al. [Mendoza et al., 2009] classify a entire lesion based on the model-based approach proposed by Serrano and Acha [Serrano and Acha, 2009]. But the main weakness of this work is the use of a low number of images from a private database.

The local features represent individual or grouped characteristics that appear in the lesion. The global features allow a quick preliminary categorization of a given pigmented skin lesion prior to more detailed assessment, and they are presented as arrangements of textured patterns covering most of the lesion. The main global patterns are: Reticular pattern, Globular pattern, Cobblestone pattern, Homogeneous pattern, Parallel pattern, Starburst pattern and Multicomponent pattern.

The main aim of this chapter is the classification of a entire pigmented lesion into Reticular pattern, Globular pattern or Homogeneous pattern by texture analysis once the lesion is segmented with the method proposed in Chapter 3. There are different reasons behind this decision instead to address the classification of the seven patterns (Reticular, Globular, Cobblestone, Homogeneous, Parallel, Starburst and Multicomponent). Globules are also predominant in the Cobblestone pattern, however they are larger and more closely aggregated than in Globular pattern, for what can be considered a special case of Globular pattern. Consequently, in our database, images belonging to Cobblestone pattern have been included in the Globular class. Regarding Parallel pattern, its automatic detection does not have a significant interest for the clinical community because lesions with this pattern are only located in palm or sole. Starburst pattern is characterized by the presence of pigmented streaks at the edge of a given lesion. As our objective is the texture analysis of an entire lesion, this type of lesion escapes from our study. Multicomponent pattern is a combination of three or more distinctive dermoscopic structures within a lesion. Therefore, the identification of different patterns, such as globular, homogeneous and reticular, implies the Multicomponent pattern identification. Nevertheless, in this chapter an unique label is assigned to the whole lesion and thus, multicomponent pattern detection is left as future work.

In this work we propose to identify the global pattern that a lesion presents by modelling in different ways. First, an image is modelled as a Markov random field in $L^*a^*b^*$ colour space to obtain texture features. In turn, these texture features are supposed to follow

different models: Gaussian model, Gaussian mixture model and a bag-of-visual words histogram model. Different distance metrics between Gaussian and Gaussian mixture distributions and between histograms are analysed. A k-Nearest Neighbour algorithm based on these distance metrics is then applied, assigning to the test image the global pattern of the closest training image.

An image database extracted from the Interactive Atlas of Dermoscopy [Argenziano et al., 2000] is used for evaluation. The results of the proposed methods are compared with the method proposed in [Mendoza et al., 2009] with our database.

The rest of the chapter is organized as follows: a review of how a textured image is modelled as a Markov random field is found in Section 5.2; in Section 5.3 the proposed classification methods are detailed; in Section 5.4 a review of the method proposed in [Serrano and Acha, 2009] is presented and the results are presented in Section 5.6. Finally, a discussion is presented.

5.2 Markov Random Field model

Models based on Markov random fields (MRF) have wide acceptance for solving texture analysis problems [Cross and Jain, 1983] [Tuceryan and Jain., 1998]. They are able to capture the local (spatial) contextual information in an image. These models assume that the intensity at each pixel in the image depends on the intensities of the neighbouring pixels [Tuceryan and Jain., 1998].

As suggested Xia et al. [Xia et al., 2006], in this thesis a finite symmetric conditional Markov (FSCM) [Kashyap and Chellappa, 1983] model characterizes the observed image to obtain texture features. These features are proposed as the basis of the different classification methods of global patterns in dermoscopic images proposed in the following Section.

The MRF model is detailed as follows: an image is considered as a random field G, defined on a $W \times H$ rectangular lattice $S = \{(i, j) : 1 \leq i \leq W, 1 \leq j \leq H\}$, which is indexed by the coordinate (i, j). The gray-scale values are represented by $G = \{G_s = g_s : s \in S\}$, where $s = (s_i, s_j)$ denotes a specific site. However, in this work, as Serrano and Acha [Serrano and Acha, 2009] proposed, the random variable G_s represents a colour pixel in the $L^*a^*b^*$ colour space instead of gray-scale values with range [0 255]. Let an observed patch $g = \{g_s : s \in S\}$ be an instance of G, defined in a square $N \times N$ center on each site s. It can be described by a finite symmetric conditional model (FSCM) [Kashyap and Chellappa, 1983] as follows:

$$g_s = \mu_s + \sum_{t \in \eta_g} \theta_{s,t} \left[(g_{s+t} - \mu_{s+t}) + (g_{s-t} - \mu_{s+t}) \right] + e_s \tag{5.1}$$

where $\eta_g = \{t_1, t_2, t_3, t_4\} = \{(0,1), (1,0), (1,1), (-1,1)\}$ is the set of shift vectors corresponding to the second order neighbourhood system, μ_s is the mean of the colour pixels in the patch centred in site s, $\{\theta_{s,t} : t \in \eta_g\}$ is the set of correlation coefficients associated with the set of translations from every site s, and $\{e_s\}$ is a stationary Gaussian noise sequence with variance σ_s^2.

Based on this FSCM model, a texture feature vector is defined as:

$$f = (\mu_s, \widehat{\sigma}_s^2, \widehat{\theta}_{s,t} : t \in \eta_g) \tag{5.2}$$

where μ_s is the mean of the colour pixels of the patch under study, $\widehat{\sigma}_s^2$ is the estimation of the noise variance, and the other four components, $\widehat{\theta}_{s,t}$, are the estimation of the correlation coefficients. In this work these features are computed from the $L^*a^*b^*$ colour space. Serrano and Acha [Serrano and Acha, 2009] analysed the use of this colour space in two cases. In the first case, the six parameters obtained from L^* component were supposed to be independent from the 12 parameters calculated from a^* and b^* components, obtaining a vector f decomposed into two parameters vectors f_L and f_{ab}. And in the second case, the feature vector was formed by 18 components, assuming independence between L^*, a^* and b^*. Their results showed that the second assumption outperformed the first one. This is reason why in this work we work with the following feature vector of 18 components:

$$f = (\mu_L, \mu_a, \mu_b, \widehat{\sigma}_L^2, \widehat{\sigma}_a^2, \widehat{\sigma}_b^2, \widehat{\theta}_{L,t}, \widehat{\theta}_{a,t}, \widehat{\theta}_{b,t} : t \in \eta_g) \tag{5.3}$$

The parameters of the FSCM model are estimated by the least-squares estimation method proposed by Manjunath and Chellappa [Manjunath and Chellappa, 1991]. Consider a region $N \times N$ (patch) containing a single texture. Let Ω be the set of all the sites belonging to the patch under consideration and Ω_I be the interior of the region of Ω, i.e. $\Omega_I = \Omega - \Omega_B$, $\Omega_B = \{s = (s_i, s_j), s \in \Omega$ and $s \pm t \notin \Omega$ for at least some $t \in \eta_g \}$:

$$\widehat{\theta} = \left[\sum_{\Omega_I} Q_s Q_s{}^T\right]^{-1} \left[\sum_{\Omega_I} Q_s g_s\right] \tag{5.4}$$

$$\widehat{\sigma}^2 = \frac{1}{N^2} \sum_{\Omega_I} \left[g_s - \widehat{\theta}^T Q_s\right]^2 \tag{5.5}$$

$$\mu = \frac{1}{N^2} \sum_{\Omega} g_s \tag{5.6}$$

where Qs is defined by $Qs = [g_{s+t} + g_{s-t} : t \in \eta_g]$ with dimensions $(4 \times 1 \times 3)$ because g_s is a 3D colour pixel. Because we are processing colour images, the dimensions of the parameters are $(4 \times 1 \times 3)$ for $\widehat{\theta}$, (1×3) for $\widehat{\sigma}^2$ and (1×3) for μ.

5.3 Proposed Model-based Classification methods

In this section the proposed model-based classification methods are detailed. The aim is the classification of a whole lesion, not only of a sample or patch of it.

It is important to note that we work with two sets of images; the first is composed by complete lesions and the second set by individual patches belonging to the lesions, each patch from a different image. This second set will be used only as training set in some methods, never as test set. In order to analyse a whole lesion, this is divided into overlapping samples. The size of each patch is fixed to 81×81 pixels, as well as each

sample in the lesion. A displacement equal to 9 rows or/and 9 columns on the lesion is applied to obtain the following sample (see Fig. 5.1 (d) (e)). In Fig. 5.1 individual patches of the three global patterns under study as well as an example of a lesion divided into overlapping samples can be seen. In Fig. 5.1 (d) and (e) the patches are obtained after a displacement of 27 rows or/and 27 columns to a right visualization. Only the patches without background or with a background area of up to 10 % the sample area are taken into account.

5.3.1 Gaussian model-based method

This approach is based on the assumption that the MRF features of the patches or samples constituting a test lesion (F_j) follow a multivariate Gaussian distribution model ($F_j \sim N(M_j, \Sigma_j)$) with mean M_j and covariance matrix Σ_j:

$$F_j \sim N(M_j, \Sigma_j) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_j|}} \exp\left(-\frac{1}{2}(f - M_j)^T \Sigma_j^{-1} (f - M_j)\right) \tag{5.7}$$

where n is the dimension of the feature vector (n=18).

Apart from this assumption, two different scenarios regarding to the training sets have been considered:

a) If individual patches are considered as the training set, MRF features of the patches belonging to each class (F_λ) are supposed to follow a multivariate Gaussian distribution with mean M_λ and covariance matrix Σ_λ.

$$F_\lambda \sim N(M_\lambda, \Sigma_\lambda) \tag{5.8}$$

b) If full lesions constitute the training set, MRF features of the patches within each training lesion (F_i) are supposed to follow a multivariate Gaussian distribution.

$$F_i \sim N(M_i, \Sigma_i) \tag{5.9}$$

Different distance metrics are used in order to compare the multivariate Gaussian distributions of the test lesion and those from the training sets. Symmetric Kullback Leibler distance [Kullback, 1997], Bhattacharyya distance [Bhattacharyya., 1943] and Frechet distance [Dowson and Landau, 1982], which is the closed form solution of the earth movers distance (EMD) in the case of two Gaussian distributions, are analysed.

The closed form expression for the symmetric KL divergence between two multivariate Gaussian distributions can be written as [Abou-Moustafa et al., 2010]:

$$SKL(F_i, F_j) = \frac{1}{2} u^T (\Sigma_i^{-1} + \Sigma_j^{-1}) u + \frac{1}{2}\left(tr(\Sigma_i^{-1}\Sigma_j) + tr(\Sigma_j^{-1}\Sigma_i) - 2n\right) \tag{5.10}$$

where $u = (M_i - M_j)$.

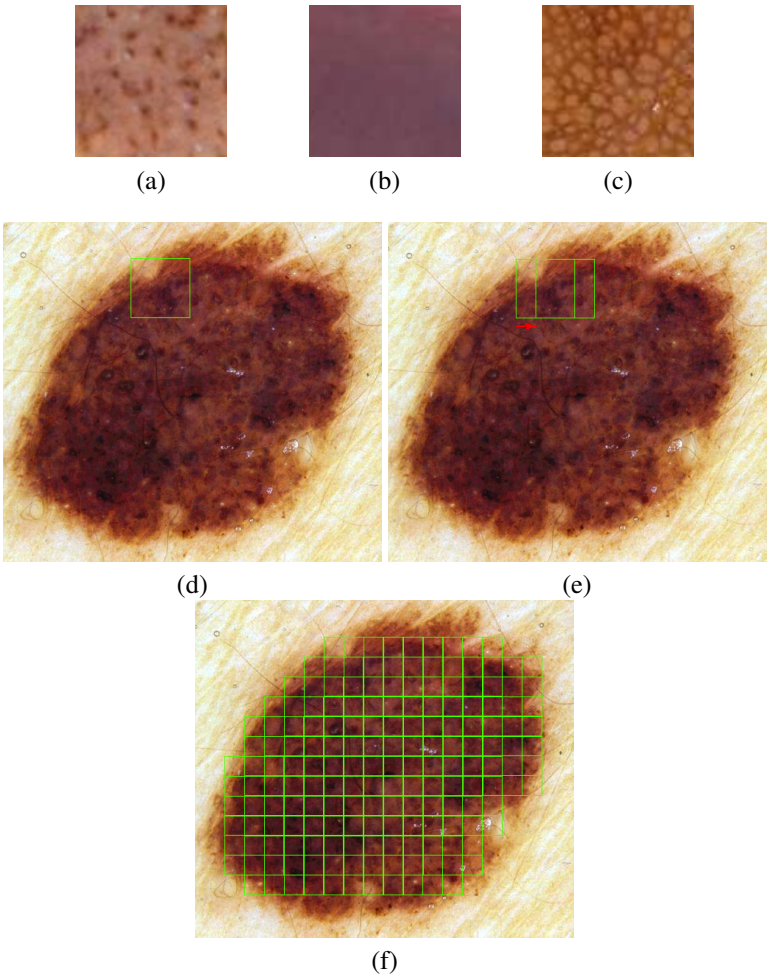


Figure 5.1 Examples of the two image sets used. a) b) c) first set: individual patches. 81×81 dermoscopic individual patches belonging to a) globular pattern, b) homogeneous pattern and c) reticular pattern. d) e) f) Second image set: complete lesions. d) 81×81 sample extracted from the whole lesion. e) A displacement equal to 27 rows or/and 27 columns is applied to obtain the following sample. f) Overlapping samples to analyse the whole lesion..

Bhattacharyya distance between two Gaussian kernels, is defined as:

$$B(F_i,F_j) = \frac{1}{8}u^T \left(\frac{\Sigma_i+\Sigma_j}{2}\right)^{-1} u + \frac{1}{2}ln\left[\frac{|\frac{\Sigma_i+\Sigma_j}{2}|}{\sqrt{|\Sigma_i||\Sigma_j|}}\right] \tag{5.11}$$

Frechet distance is computed as:

$$Fr(F_i,F_j) = \sqrt{\left(u^T u + tr(\Sigma_i + \Sigma_j - 2\sqrt{\Sigma_i \Sigma_j})\right)} \tag{5.12}$$

The k-Nearest Neighbour algorithm (KNN) with the aforementioned distances has been applied for the final classification. In the first scenario, a test image is identified with the pattern closest to it. In the second case, a KNN approach is applied so that the test image is assigned to the class of the training image closest to it.

5.3.2 Gaussian mixture model-based methods

According to Sfikas et al. [Sfikas et al., 2005], in the context of image retrieval, it is advantageous to model the feature data using parametric probability density function models, such us Gaussian mixture models (GMM). In this approach MRF features extracted from patches constituting a test lesion are supposed to follow a Gaussian mixture model. This model represents a probability density function (PDF) as:

$$p(f) = \sum_{j=1}^{K} \pi_j N(M_j, \Sigma_j) \tag{5.13}$$

where K stands for the number of Gaussian kernels mixed, M_j and Σ_j are the mean vectors and the covariance matrices of Gaussian kernel j and π_j are the mixing weights. These parameters and weights are estimated iteratively from the input MRF features using the expectation-maximization (EM) algorithm [Mclachlan and Peel, 2000]. In this approach the number of the Gaussian components for every Gaussian mixture model (GMM) was empirically chosen to be three.

Based on this assumption, other two suppositions, regarding to the training sets, are considered, similarly to Section 5.3.1:

a) The MRF features of the individual training patches belonging to each class follow a Gaussian mixture distribution.

$$p_\lambda(f) = \sum_{i_\lambda=1}^{L} \pi_{i_\lambda} N(M_{i_\lambda}, \Sigma_{i_\lambda}) \tag{5.14}$$

where λ represents each global pattern, L stands for the number of Gaussian kernels mixed for each pattern, M_{i_λ} and Σ_{i_λ} are the mean vectors and the covariance matrices of Gaussian kernel i_λ and π_{i_λ} are the mixing weights.

b) The training set now consists of full lesions that are supposed to follow a Gaussian mixture distribution.

$$p'(f) = \sum_{i=1}^{L} \pi'_i N(M'_i, \Sigma'_i) \tag{5.15}$$

As in the previous approach (see Section 5.3.1), the idea is to compare the Gaussian mixture model of a test lesion with the mixture distribution corresponding to the training sets. To this purpose different distance metrics between Gaussian mixture models are used: the symmetric Kullback Leibler divergence [Sfikas et al., 2005], the Bhattacharyya-based distance metric [Sfikas et al., 2005], the earth movers distance (EMD) [Greenspan et al., 2004] and a distance metric proposed by Sfikas et al. [Sfikas et al., 2005]. This metric is computed as:

The symmetric Kullback Leibler divergence for GMMs can be computed as:

$$SKL(p,p') = |\frac{1}{2N}\sum_{f\sim p} lnp(f) - \frac{1}{2N}\sum_{f\sim p} lnp'(f) + \\ + \frac{1}{2N}\sum_{f\sim p'} lnp(f) - \frac{1}{2N}\sum_{f\sim p'} lnp'(f)| \tag{5.16}$$

where N is the minimum number of data samples (features vectors from the patches) and $p(f)$ and $p'(f)$ are estimated iteratively with expectation maximization (EM) [Sfikas et al., 2005].

The Bhattacharyya-based distance metric is computed as:

$$BhGMM(p,p') = \sum_{i=1}^{L}\sum_{j=1}^{K} \pi_j \pi'_i B(p_j, p'_i) \tag{5.17}$$

where p, p' are Gaussian mixture models consisting of K and L kernels respectively. p_j, p'_i denote the kernel parameters and π_j, π'_i are the mixing weights. B denotes the Bhattacharyya distance between two Gaussian kernels (Equation (5.11)).

The earth movers distance (EMD) for GMMs is computed as:

$$EMD(p,p') = \frac{\sum_{i=1}^{L}\sum_{j=1}^{K} f_{ij} Fr(p_j, p'_i)}{\sum_{i=1}^{L}\sum_{j=1}^{K} f_{ij}} \tag{5.18}$$

$$\sum_{i=1}^{L}\sum_{j=1}^{K} f_{ij} = min\left(\sum_{j=1}^{K}\pi_j, \sum_{i=1}^{L}\pi'_i\right) \tag{5.19}$$

where $Fr(p_j, p'_i)$ is the Frechet distance between p_j and p'_i (Equation (5.12)).

The distance metric proposed by Sfikas et al. [Sfikas et al., 2005] to compare Gaussian Mixture is computed as:

$$C2(p,p') = -log\left[\frac{2\sum_{i,j}\pi_i\pi'_j\sqrt{\frac{|V_{ij}|}{e^{k_{ij}}|\Sigma_i||\Sigma'_j|}}}{\sum_{i,j}\left\{\pi_i\pi_j\sqrt{\frac{|V_{ij}|}{e^{k_{ij}}|\Sigma_i||\Sigma_j|}}\right\} + \sum_{i,j}\left\{\pi'_i\pi'_j\sqrt{\frac{|V_{ij}|}{e^{k_{ij}}|\Sigma'_i||\Sigma'_j|}}\right\}}\right] \tag{5.20}$$

$$V_{ij} = \left(\Sigma_i^{-1} + \Sigma_j'^{-1}\right)^{-1} \tag{5.21}$$

$$k_{ij} = M_i^T \Sigma_i^{-1}(M_i - M_j') + M_j'^T \Sigma_j'^{-1}(M_j' - M_i) \tag{5.22}$$

π, π' the mixing weights, i and j are indexes on the Gaussian kernels, and M, Σ and M', Σ' are mean and covariance matrices for the kernels of Gaussian mixture $p(f)$ and $p'(f)$ respectively.

As in the previous Section, a k-Nearest Neighbor algorithm is applied. In the first case, a test image is identified with the closest pattern according to the four distances proposed above and in the second case the test image is assigned to the pattern of the closest training image.

5.3.3 Bag of features

The last approach is based on the representation of an image as a bag of features (BoF). This approach finds its origin, on the one hand, in the texture recognition by textons [Julesz, 1981] [Varma and Zisserman, 2003] (basic elements of texture) and, on the other hand, in the bag of words scheme used for text categorization and text retrieval [Lewis, 1998]. The idea is to model an image as a frequency histogram of visual words (bag of features). These visual words are built from the quantization of descriptors (in our case the descriptors are MRF features) of local patches sampled from the training set. This quantization is usually carried out by a clustering algorithm such as k-means. The centroid of each cluster represents a visual word. The set of visual words forms a codebook.

The bag of features representation is widely used in computer vision for classification of natural scenes [Csurka et al., 2004] or biomedical images such as histology images [Cruz-Roa et al., 2011] or different tissues [Wang et al., 2013]. However, to the best to our knowledge, only a few works can be found in the literature that use bag of features applied to pigmented lesions. Situ et al. [Situ et al., 2008] [Situ et al., 2010] applied this approach to classify lesions between malignant and benign. Recently, Barata et al. [Barata et al., 2013] also proposed a binary classification between malignant and benign lesions using bag of features focusing in different strategies for the extraction of interest regions from which features are extracted. Sadeghi et al [Sadeghi et al., 2012a] used an approach based on textons and applied to filter bank responses for each pixel to identify different global patterns, however only individual patches are classified, not whole image.

In our case, for each global pattern, the MRF features of the patches of the training lesions form a n-dimensional space. Specifically, we have *No patches per image* $\times$ *No training images per class* n-dimensional vectors per class located in this space ($n =$ 18). These n-dimensional vectors belonging to the same class are clustered by K-means algorithm, obtaining K centroids or visual words for each pattern. These 3*K (3 patterns under study) visual words form the codebook. Then, all the patches from all training images are reassigned to the closest visual word of the codebook. An histogram with the frequency of occurrences of each centroid or visual word is formed for each training image.

In Fig. 5.2 an overview of the proposed BoF approach to image classification is shown.

In the classification step, overlapping patches are extracted from a new test image and a n-dimensional vector with MRF features is estimated from each patch. Each n-dimensional vector is assigned to the nearest centroid of the codebook, forming a histogram of frequencies of clusters (bag of features). Finally, a classifier is applied to identify the pattern that this model belongs to, regarding to the models from the training image (see Fig. 5.2).

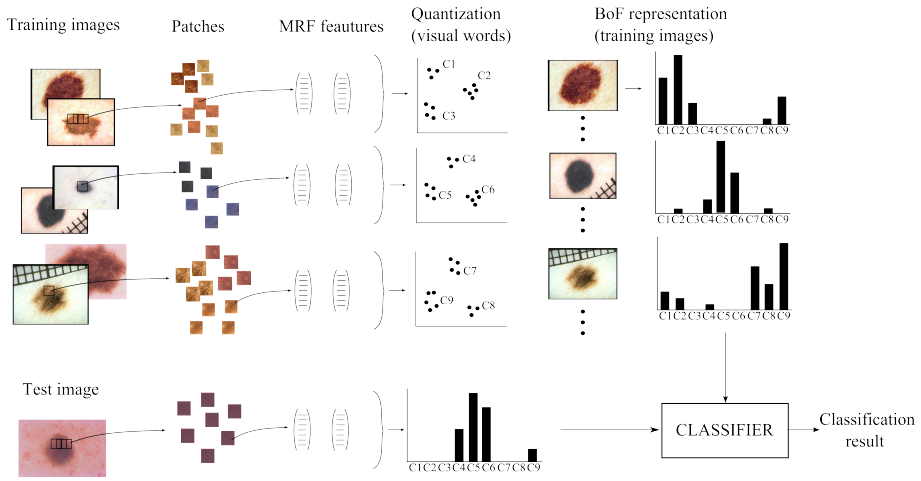


Figure 5.2 Overview of the BoF approach to image classification.

A K-Nearest Neighbor algorithm (KNN) with different histogram dissimilarity measures is proposed as classifier. Five common dissimilarity measures between two histograms $H = h_i$ and $K = k_i$ (i represents bins) are used [Rubner et al., 2000]: earth movers distance (EMD), χ^2 Statistics, histogram intersection, Kullback-Leibler divergence and Kolmogorov-Smirnov distance. Below their computation is indicated.

Earth Movers Distance: Let $P = (p_1, w_{p_1}), ..., (p_m, w_{p_m})$ be the first histogram with m bins, where p_i is the bin representative and w_{p_i} is the weight of the bin; $Q = (q_1, w_{q_1}), ..., (q_m, w_{q_m})$ the second histogram with n bins; d_{ij} is the ground distance between p_i and q_i defined as norm L_1; and f_{ij} the flow between p_i and q_i subject to the following constraints:

$$\begin{aligned} f_{ij} &\geq 0, \quad 1 \leq i \leq m, 1 \leq j \leq n \\ \sum_{j=1}^{n} f_{ij} &\leq w_{p_i}, \quad 1 \leq i \leq m \\ \sum_{i=1}^{m} f_{ij} &\leq w_{q_j}, \quad 1 \leq i \leq m \\ \sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij} &= min\left(\sum_{i=1}^{m} w_{p_i}, \sum_{j=1}^{n} w_{q_j}\right) \end{aligned} \tag{5.23}$$

$$EMD(P,Q) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} d_{ij} f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}} \tag{5.24}$$

χ^2 Statistics:

$$d_{\chi^2}(H,K) = \sum_i \frac{(h_i - m_i)^2}{m_i} \tag{5.25}$$

$$m_i = \frac{h_i + k_i}{2} \tag{5.26}$$

Histogram Intersection:

$$d_{\cap}(H,K) = 1 - \frac{\sum_i min(h_i,k_i)}{\sum_i k_i} \tag{5.27}$$

Kullback-Leibler Divergence:

$$d_{KL}(H,K) = \sum_i h_i log \frac{h_i}{k_i} \tag{5.28}$$

Kolmogorov-Smirnov distance:

$$d_{KS}(H,K) = max_i(|\widehat{h_i} - \widehat{k_i}|) \tag{5.29}$$

where $\widehat{h_i} = \sum_{j \le i} h_j$ is the cumulative histogram of h_i, and similarity for k_i.

According to [Csurka et al., 2004], on natural image classification the larger the codebook size the better. However, Tomassi et al. [Tommasi et al., 2008] found that the size of the codebook is not a significant aspect in a medical image classification task. In this work different codebook sizes are evaluated (see Fig 5.6 of Section 5.6).

5.4 Classification method by computation of posterior probability

The proposed methods are compared with the method proposed in [Mendoza et al., 2009], which it is based on [Serrano and Acha, 2009], applied to our database. In this Section, a review of the technique proposed in [Serrano and Acha, 2009] is carried out. First, it is important to note that in this method both the training set and the test set are composed by individual patches. However, [Mendoza et al., 2009] considered a lesion divided by overlapping patches, and the final classification of whole lesion was made by polling process.

The method is based on the MAP-MRF framework [Dubes and Jain, 1989] that suggested that the optimal pattern under a feature set f can be obtained by maximizing the posterior probability. The following assumption is introduced to calculate this probability: the features of the training patches F of each class, follows a multivariate Gaussian distribution $N(M_\lambda, \Sigma_\lambda)$ with mean vector M_λ and covariance matrix Σ_λ, corresponding to the pattern λ belongs to. λ is an instance of a random variable Λ, taking values from a finite set $R = \{1,2,3\}$. MRF features of the training patches belonging to each class form

a Gaussian distribution as:

$$N(M_\lambda, \Sigma_\lambda) = \\ = \frac{1}{\sqrt{(2\pi)^n |\Sigma_\lambda|}} \exp\left(-\frac{1}{2}(f - M_\lambda)^T \Sigma_\lambda^{-1} (f - M_\lambda)\right) \tag{5.30}$$

where n is the dimension of the feature vector ($n = 18$).

Therefore, in order to classify each sample of the test lesion with feature vector f into the pattern it belongs to, the *maximum a posteriori* (MAP) criterion is applied together with the assumption that the three possible global patterns (globular, homogeneous and reticular) are equally probable, what results in the *maximum likelihood* (ML) criterion. Then:

$$\hat{\lambda} = \arg\max_{\lambda \in \Lambda} P(F = f | \lambda) \tag{5.31}$$

where f is the vector of MRF features for the patch to be classified. This ML problem can then be solved by minimizing the following energy:

$$\hat{\lambda} = \min_{\lambda \in \Lambda} E(f, \lambda) = \\ = \min_{\lambda \in \Lambda} \{((f - M_\lambda)^T \Sigma_\lambda^{-1} (f - M_\lambda) + \ln((2\pi)^n |\Sigma_\lambda|)\} \tag{5.32}$$

In [Mendoza et al., 2009], each patch of the lesion is classified in this way and the whole lesion is assigned to the global pattern most voted.

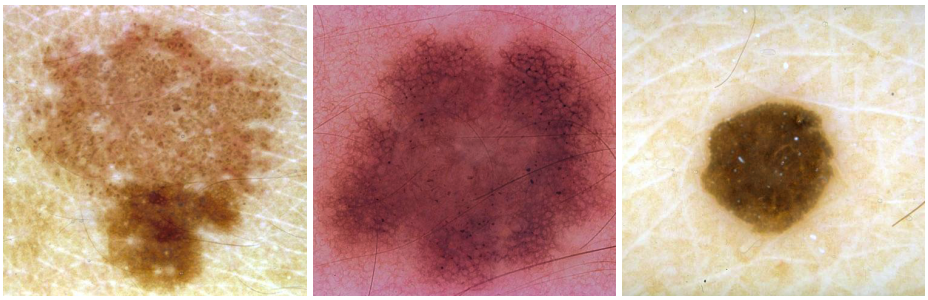
5.5 Image Database

The image database used in this work is formed by 30 images of each type of pattern, a total of 90 images. As it has already been mentioned, globules are predominant in Globular and Cobblestone pattern, however, for the second case, they are larger and more closely aggregated than in Globular pattern, for what can be considered a special case of Globular pattern. 8 images of the 30 categorized as globular pattern, belong to Cobblestone pattern.

All images were extracted from the Interactive Atlas of Dermoscopy, published by Edra Medical Publishing New Media [Argenziano et al., 2000], which is a multimedia project for medical education with images of pigmented skin lesions from different centres and hospitals.

Considering that the global pattern is determined by the dermatoscopic feature predominant in the lesion, its automated classification becomes hard due to the possible presence of different patterns in the same lesion. An example can be seen in Fig. 5.3 (a), where the lesion was classified as globular, however, the inferior part of the lesion is reticular. Fig. 5.3 (b) shows a lesion whose classification was reticular, but also presents globules and homogeneous areas.

Besides this intrinsic difficulty, the images from this Atlas of Dermoscopy present two difficulties for their automatic classification: intra-class variability, lesions belonging to the same global pattern with very different appearance, and inter-class similarity, lesions



(a) (b) (c)

Figure 5.3 Examples of images from the database. (a) Image classified as globular pattern. (b) Image classified as reticular pattern. (c) Image classified as homogeneous pattern..

belonging to the different global pattern with certain similar appearance.

Moreover, the virtual Consensus Net Meeting on Dermoscopy (CNMD) [Argenziano et al., 2003] was organized to investigate the interobserver and intraobserver reproducibility and validity of the various features and diagnostic algorithm. The reproducibility was assessed according to the method of Fleiss et al. [Fleiss, 1981] to calculate the κ statistics. Its interpretation is: a value of 1.0 indicates perfect agreement and values less than 0.4 are poor. 128 dermoscopic images of pigmented skin lesions were selected to include in this study. The interobserver reproducibility was computed among 40 observers. To test for intraobserver agreement, 20 lesions were randomly selected and included for re-examination. The results presented for diagnostic of Global patterns were $\kappa = 0.43$ for interobserver agreement and $\kappa = 0.55$ for intraobserver agreement. These results show the difficulty in the diagnosis of global patterns.

In this work, the reference truth has taken according to the diagnosis presented in the Interactive Atlas of Dermoscopy [Argenziano et al., 2000].

5.6 Evaluation and results

To evaluate the performance of the proposed methods a k-fold cross validation is used. There are two goals in cross-validation [Refaeilzadeh et al., 2009]: To estimate performance of the learned model from available data using one algorithm. In other words, to gauge the generalization of an algorithm. And to compare the performance of different algorithms and find out the best algorithm for the available data.

In k-fold cross-validation, the image database is randomly partitioned into k equal size groups or splits. Of the k groups, a single groups is used as the validation data for testing the model, and the remaining k-1 groups are used as training data. The cross-validation process is then repeated k times (folds), with each of the k groups used exactly once as the validation data. The k results from the folds then are averaged to produce a single estimation. The advantage of this method is that all images are used for both training and

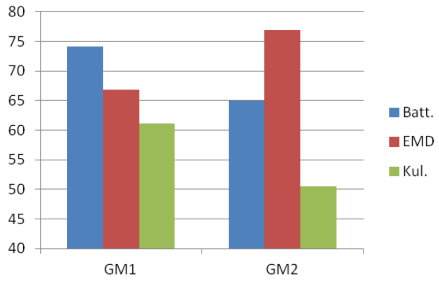


Figure 5.4 Performance of the two methods proposed in Section 5.3.1 (GM1 and GM2) when using different distance metrics between probability density functions: Bhattacharyya distance (Batt.), earth movers distance (EMD) and Kullback Leibler divergence(Kul.). y-axis shows the accuracy of the classification methods. .

validation, and each image is used for validation exactly once.

However, due to the variability of our database, as mentioned in Section 5.5, the choice of the k splits has a high influence in the results. A full cross-validation, i.e. performing all-possible ways of partitioning, would give an accurate estimation, but it is computationally too expensive. Therefore, repeating k-fold cross-validation r-times using different random splits for each run, provides a good Monte-Carlo estimate of the full cross-validation [Kohavi, 1995]. In this work, we used a 20-times 3-fold cross-validation.

For all the methods, a comparative study of the proposed metrics was performed.

The performance of the three distance metrics proposed in the two methods based on multivariate Gaussian model (GM1 when the training set are the individual patches and GM2 when the training set are the entire lesions), is presented in Fig 5.4. It can be seen that a different distance outperforms the rest for each method. For the first method, Bhattacharyya is superior than the rest, however, in the second scenario EMD offers better performance. Similarly, for the methods proposed in Section 5.3.2, the distance proposed in [Sfikas et al., 2005] outperforms the rest in the first scenario whereas in the second case EMD outperforms the rest of distances (see Fig. 5.5).

Regarding the Bag of Features approach (Section 5.3.3), Fig. 5.6 shows the performance for the different histogram dissimilarity measures: earth movers distance (EMD),χ^2 statistic (χ^2), histogram intersection (Hist.), Kolmogorov-Smirnov distance (Kol.) and Kullback-Leibler divergence(Kul.). They have been evaluated with different number of centroids or visual words. In view of the results in Fig. 5.6, it seems that the number of visual words does not significantly influence the success rate. However, χ^2 distance using 20 centroids per class (60 visual words in total) achieved the best result.

Table 5.1 shows the accuracy of the classification for all proposed methods. Only those with the highest classification rate obtained in Fig. 5.4, Fig. 5.5 and Fig. 5.6 are presented. In addition, the individual accuracy in the identification of each global pattern is shown. In general, homogeneous pattern is identified with a success rate of over 90% in all cases, decreasing this rate for globular and reticular pattern identification. It can be conclude

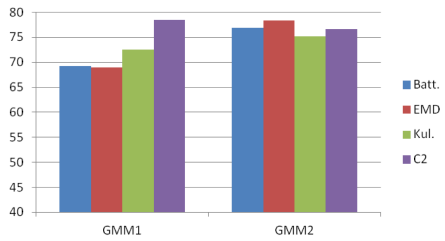


Figure 5.5 Performance of the two methods proposed in Section 5.3.2 (GMM1 and GMM2) when using different distance metrics between Gaussian mixture models: Bhattacharyya-based (Batt.), earth movers distance (EMD), Kullback Leibler divergence(Kul.) and a distance proposed in [Sfikas et al., 2005] (C2). y-axis shows the accuracy of the classification methods. .

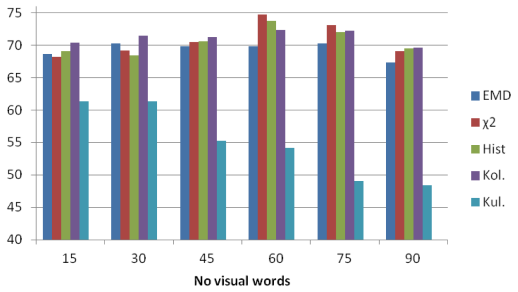


Figure 5.6 Performance of the Bag of Features approach when different codebook sizes are used. KNN on different histogram dissimilarity measures: earth movers distance (EMD), χ^2 statistic (χ^2), histogram intersection (Hist.), Kolmogorov-Smirnov distance (Kol.) and Kullback-Leibler divergence(Kul.) y-axis shows the accuracy of the classification methods. .

that Gaussian mixture model-based methods outperform the rest in average.

The method proposed in [Mendoza et al., 2009] has been included in the evaluation (see Table 5.1). The results show that the proposal has significantly better performance.

A further evaluation was performed. Once three global patterns were correctly classified, we include the multicomponent pattern in the study, which is characterized by the presence of three or more patterns. 30 images of melanomas with multicomponent pattern were chosen randomly from the Interactive Atlas of Dermoscopy [Argenziano et al., 2000]. Examples of melanoma are shown in Fig. 5.7. The evaluation has been performed using the method based on Gaussian mixture. As the multicomponent pattern is a combination of various patterns the extraction of patches here is not convenient since a sample of the lesion cannot contain the different textures from all dermoscopic features that the lesion presents. Therefore, we only applied the GMM2 method assuming that training set is constituted by lesions. Table 5.2 presents the results. Although the success rate in

Table 5.1 Results of classification for the proposed methods compared with the method proposed in [Mendoza et al., 2009]. Bold text indicate the better success rates obtained. .

Method	Globular	Homogeneous	Reticular	Average
GM1 (Batt.)	**70.50** %	90.67%	61.33%	74.16%
GM2 (EMD)	62.50%	99.66%	68.67%	76.94%
GMM1 (C2)	62.00%	98.00%	**75.33** %	**78.44**%
GMM2 (EMD)	66.50%	**99.67**%	69.00%	78.38%
BoF (χ^2)	66.50%	97.67%	60.00%	74.72%
E [Mendoza et al., 2009]	52.83%	74.83%	53.83%	60.50%

(a) (b) (c)

Figure 5.7 Examples of melanomas images with multicomponent pattern..

Table 5.2 Results of classification for the GMM2 methods when lesions with multicomponent pattern are included in the study. .

Method	Glob.	Homog.	Retic.	Multic.	Average
GMM2 (EMD)	64.33%	95.83%	67.00%	64.5%	72.91%

average has decreases in 5.4% a good classification is still reached. What demonstrates the good performance of modelling.

5.7 Summary and conclusions

In this chapter classification methods of global dermoscopic patterns have been proposed. The aim is to classify each lesion with a global pattern. This unique-label classification is motivated by the fact that a lesion is characterized by a global pattern and by one or more local patterns. The majority of the classification approaches in the literature are based on a feature extraction step followed by a classifier whose inputs are the features extracted. On the contrary, we propose techniques based on modelling in different senses. First, an image is modelled by a Markov random field on the $L^*a^*b^*$ colour space. The

estimated parameters of this model are treated as features. And then, these features within a lesion are supposed to follow three different models. In the first one, it is supposed that a lesion follows a multivariate Gaussian distribution. The idea is to measure distances between Gaussian models (GM) and the to apply a KNN algorithm. The second approach proposed remains the same idea but supposing Gaussian mixture models (GMMs). As in the previous case different distance metrics between GMMs are analysed. The third model-based classification technique is a Bag of Features approach, where a image is modelled by a frequencies histogram of visual words. In this case, different distances between histograms have been studied.

For the evaluation a database from the Interactive Atlas of Dermoscopy [Argenziano et al., 2000]. A comparative study between the classification success rates obtained by all methods is presented. In view the results, it can be concluded that the methods based on the assumption that MRF features, which characterize a lesion and consequently a global pattern, follow a Gaussian Mixture model outperform the rest, obtaining a 78,44% on average when three global patterns (reticular, globular/cobblestone, homogeneous) are classified. A further evaluation was performed. 30 images of melanomas with multi-component pattern were included achieving for this case a 72,91%. This result shows the potential of this system for early melanoma diagnosis.

Moreover, the proposed model-based methods have been compared with the method of a previous work proposed by Serrano and Acha [Serrano and Acha, 2009], in which individual patches were classified into five global patterns, and later applied to whole lesion by Mendoza et al. [Mendoza et al., 2009]. The results evidence the superiority of the methods here proposed.

5.8 Discussion

The first novelty presented in this chapter is that features within a lesion are modelled. In other words, a multidimensional histogram is formed with the features within a lesion and this histogram is modelled with a particular density model. Then, classification is perform via comparison of histogram or density models, with specific dissimilarity functions.

Serrano et al. [Serrano and Acha, 2009] used MRF features in $L^*a^*b^*$ to model global patterns in patches. But then they modelled as a Gaussian distribution of MRF features within a class in a database. Differently, in this work, for the first time, the distribution of MRF features within a lesion is analysed.

The second novelty consist in applying the concept of bag of visual words to MRF features of pigmented lesions. It is a very convenient approach to analyse very sparse histogram in a n-dimensional space. Previously, in [Sadeghi et al., 2012a] the concept of textons to the problem of global pattern classification was applied. Nevertheless, differently from our approach, they applied the textons concept to pixels within a patch and they used as features the output of filter banks. In our proposal bag of features approach is applied to MRF features of the different patches within a lesion.

Finally, it should be outlined that no previous attempts of global pattern model-based classification of full lesions into global patterns can be found in the literature.

CHAPTER 6

6 Summary and conclusions

¿Tomas algo para ser feliz? Sí, decisiones

AUTOR DESCONOCIDO

This thesis presents different techniques within the image processing framework to solve two biomedical computing challenges: neuromuscular diseases classification from fluorescece miscroscopic images of muscle biopsies and pattern classification for pigmented lesions in dermoscopic images.

The first part of this thesis addresses the neuromuscular diseases classification problem. A system for diagnostic support of muscular dystrophies and neurogenic atrophies is presented thoroughly in Chapter 2. The main conclusions of this work are summarized in the following points.

The system proposed includes the main steps constituting a computer-aided diagnosis system: segmentation, feature extraction, feature selection and classification. Simple techniques used in each step have achieved a successful solution to the problem raised obtaining a high success rate in the classification.

The segmentation algorithm proposed uses mathematical morphology and a watershed transformation. A Dice index of 0.967 and a Jaccard index of 0.934, in average, are achieved.

The main novelty of the system lies in the feature extraction step. Not only features related to the characteristics that the pathologist takes into account for diagnosis but features with inherent properties that escape to the evaluation of the pathologist are extracted. In this sense, an extraction of structural information based on the representation of a biopsy as a graph, where the nodes are represented by each fibre and two nodes are connected if two fibres are adjacent, has been proposed.

Specially important is the study of the effectiveness of the two set of features proposed. It is concluded that the best set of features to distinguish between controls

and muscular dystrophies is that derived from the features that the physicians take into account in their diagnosis. This fact is justified by clear differences between the biopsies belonging to these two groups. However, the addition of new features undetectable by the human visual inspection improves the identification of biopsies with atrophic pattern. This is due to the fact that biopsies affected by atrophy present similar characteristics to control biopsies. This finding represents a breakthrough in the early diagnosis of muscular atrophies.

A severity grading study has been performed for biopsies of distrophic pattern by applying a principal component analysis. The Pearson correlation is used to study if there is a relationship between the location in the PCA graph and the affection degree of a biopsy. The high value (0.875) obtained indicate that PCA can be considered as a way to analyse the affectation degree.

In future works the study will be extended to other neuromuscular diseases. The final aim is the development of a user interface with the purpose that any neurologist, pathologist or researcher can analyse their own images using our method. This would provide a useful service to the biomedical community both for diagnosis and research activities.

The second part of this thesis deals with pattern analysis in pigmented lesions from dermoscopic images in an attempt to diagnose the malignity of these lesions.

The first step was to develop a segmentation algorithm (Chapter 3) to isolate the lesion from the background skin as a necessary step before classification. The main challenge was to achieve an algorithm to be correlated with human colour perception. The segmentation method proposed is based on a perceptually adapted gradient integrated in a level-set framework for edge detection in colour images. Different gradients based on three colour differences (CIELAB, CIE94, CIEDE2000) were tested. CIE94 based edge detector proved to have the highest correlation with what the human eye can perceive. Its application in pigmented lesion segmentation resulted to be successful despite the great variety of lesion shapes, sizes, and colours along with different skin types and textures. The main advantages in the segmentation method proposed are the application of the procedure without requiring a preprocessing to remove artefacts, the inclusion of colour information and its automatic nature. The last aspect becomes crucial if the final objective is the development of a CAD system.

Thereafter, this thesis tries to emulate pattern analysis, the technique of diagnosis of pigmented skin lesions considered superior to other algorithms. Pattern analysis seeks to identify specific patterns, which may be global and local. An extensive review of algorithmic methods found in literature that automatically detect these patterns in dermoscopic images of pigmented lesions is presented in Chapter 4. It is concluded that numerous works have focused on the extraction of local patterns, however, when dealing with the detection and/or classification of global patterns, only a few methods have been published in the literature.

Considering the previous conclusion, novel global pattern classification methods are proposed in Chapter 5. These methods are based on modelling techniques. First

a dermoscopic image is modelled by Markov random fields applied to $L^*a^*b^*$ colour space and the estimated parameters of this model are treated as features. In turn, the distribution of these features are supposed that follow different models along a lesion: a Gaussian model, a Gaussian mixture model and a bag-of-features model. The aim is to identify four global patterns: reticular, globular, cobblestone and homogeneous, whose diagnostic significance varies. The reticular pattern is found in benign acquired melanocytic nevi in general and in thin melanomas in particular, as well as in dermatofibroma. The globular and the cobblestone pattern can be found in Clark nevi, Unna nevi and seborrheic keratosis. The homogeneous pattern represents the dermoscopic hallmark of blue nevus, although it may also be present in Clark nevi, dermal nevi, metastatic melanomas and basal cell carcinomas. It is concluded that the best classification success rate is achieved by the Gaussian mixture model-based method with a 78.44% success rate in average. A further evaluation includes the categorization of multicomponent pattern, the pattern most usual in melanoma. Their promising results (72.91% of success rate) show the potential of this system for early melanoma diagnosis.

The next step towards the development of a computer-aided diagnosis system for pigmented lesions is local pattern detection, what will permit a more detailed assessment of pigmented skin lesions.

In general, this thesis presents a wide range of image processing techniques to meet different biomedical challenges. Both simple techniques that together allow the development of a complex system to achieve a successful solution to a given problem, and novel techniques that open the way to future research lines.

CHAPTER 7

7 Conclusiones

¿Tomas algo para ser feliz? Sí, decisiones

AUTOR DESCONOCIDO

Esta tesis presenta diferentes técnicas de procesamiento de imagen para aplicaciones biomédicas. En concreto, una clasificación de enfermedades neuromusculares a partir de imágenes de microscopía de fluorescencia de biopsias musculares y una clasificación de patrones en lesiones pigmentadas en imágenes demoscópicas.

La primera parte de esta tesis aborda el problema de clasificación de enfermedades neuromusculares. En el capítulo 2 se presenta un sistema para apoyo al diagnóstico de distrofias musculares y atrofias neurógenas. Las principales conclusiones de este trabajo se resumen en los siguientes puntos.

El sistema propuesto incluye los principales pasos constituyentes de un sistema de diagnóstico asistido por ordenador: segmentación, extracción de características, selección de características y clasificación. El uso de técnicas simples en cada uno de los pasos propuestos ha permitido alcanzar una solución exitosa al problema planteado.

El algoritmo de segmentación propuesto usa morfología matemática y una transformada watershed. Se consigue un valor medio del índice de Dice de 0.967 y de 0.934 para el índice de Jaccard.

La principal novedad del sistema reside en el paso de selección de características. Se extraen no solo aquellas que el patólogo tiene en cuenta para diagnosticar sino características con propiedades que se escapan de la evaluación del patólogo. En este sentido, se propone una extracción de información estructural de la biopsia basada en la representación de su estructura como un grafo, en el que los nodos representan las fibras musculares y dos nodos están conectados si dos fibras son adyacentes.

Especialmente importante es el estudio de la efectividad de los dos conjuntos de características propuestos. Se concluye que el conjunto más discriminatorio entre controles (sin enfermedad) y distrofias musculares es aquel que deriva de los descriptores que los especialistas tienen en cuentan para tomar su diagnóstico. Este hecho se justifica por las claras diferencias existentes entre las biopsias pertenecientes a estos dos grupos. Sin embargo, la adicción de nuevas características no detectables por inspección visual mejoran la identificación de las biopsias con un patrón atrófico. Esto se debe a que este tipo de biopsias presenta características similares a las biopsias control. Este hallazgo representa un gran avance en el diagnóstico precoz de las atrofias neurógenas.

Se ha llevado a cabo también un estudio de la severidad de las biopsias con patrón distrófico mediante la aplicación de un análisis de componente principales (PCA). Se ha empleado un coeficiente de correlación de Pearson para estudiar si existe alguna relación entre la localización de las biopsias en el gráfico PCA y el grado de afectación que presenta la biopsia. El alto valor obtenido (0.875) indica que PCA puede considerarse una manera de análisis del grado de afectación.

En trabajos futuros el sistema se extenderá a otras enfermedades neuromusculares. El objetivo final es el desarrollo de una interfaz de usuario en la que cualquier neurólogo, patólogo o investigador puede analizar sus propias imágenes con nuestro método. Esto proporcionaría una servicio útil a la comunidad biomédica tanto para el diagnóstico como para la investigación.

La segunda parte de esta tesis trata del análisis de patrones en lesiones pigmentadas a partir de imágenes demoscópicas en un intento de diagnosticar la malignidad de estas lesiones.

El primer paso fue el desarrollo de un algoritmo de segmentación (Capítulo 3) para aislar la lesión de la piel circundante como paso previo a la clasificación. El principal reto fue conseguir un algoritmo con alta correlación con la percepción humana del color. El método de segmentación propuesto se basa en un gradiente perceptualmente adaptado integrado en una técnica de level set para detección de bordes en imágenes en color. Se analizaron diferentes gradientes basados en tres diferencias de color (CIELAB, CIE94,CIEDE2000). El detector de bordes basado en CIE94 obtuvo la correlación más alta con lo que el ojo humano puede percibir. Su aplicación para segmentación de lesiones pigmentadas resultó exitosa a pesar de la gran variedad en forma, tamaño, textura y tipos de piel que las lesiones presentan. Las principales ventajas del método de segmentación propuesto son la aplicación del procedimiento sin la necesidad de un preprocesamiento para eliminar artefactos, la inclusión de información de color y su naturaleza automática. Éste último aspecto se convierte en crucial si el objetivo final es el desarrollo de un sistema de diagnóstico asistido por ordenador.

A partir de entonces, la tesis intenta emular el método de análisis de patrones, la técnica de diagnóstico de lesiones pigmentadas de la piel más empleada por los especialistas. Este método identifica patrones específicos, los cuales pueden ser

locales y globales. En el Capítulo 4 se presenta una extensa revisión de los métodos algorítmicos publicados en la literatura que detectan automáticamente estos patrones. Se concluye que existen numerosos trabajos que se centran en la extracción de patrones locales, sin embargo, solo unos pocos abordan el problema de la clasificación de patrones globales.

Considerando la conclusión previa, en el Capítulo 5 se proponen métodos de clasificación de patrones globales en lesiones pigmentas. Estos métodos se basan en el análisis de textura mediante técnicas de modelado. Primero una imagen demoscópica se modela mediante campos aleatorios de Markov, los parámetros estimados de este modelo se consideran características. Después, se supone que la distribución de estas características a lo largo de una lesión sigue diferentes modelos: un modelo gaussiano, un modelo de mezcla de gaussianas y un modelo de bolsa de características. El objetivo es identificar tres patrones: reticular, globular y empedrado (considerado un solo patrón) y homogéneo. El patrón reticular se encuentra en nevos melanocíticos benignos en general y en melanoma delgado en particular, así como en dermatofibroma. El patrón globular y el patrón empedrado se encuentran en el nevo de Clark, nevo de Unna y en queratosis seborreica. El patrón homogéneo representa el sello dermatoscópico del nevo azul, aunque también puede estar presente en nevos de Clark, nevos dérmicos, melanomas metastáticos y carcinomas de células basales. Se llegó a la conclusión de que el método basado en modelo de mezcla de gaussianas consigue la mejor tasa de éxito de clasificación, con un valor en promedio de un 78,44%. Además se incluye una evaluación adicional en la que se clasifica melanomas con patrón multicomponente. Sus resultados prometedores (72,91 % de tasa de éxito) muestran el potencial de este sistema para el diagnóstico precoz del melanoma.

El siguiente paso hacia el desarrollo de un sistema de diagnóstico asistido por ordenador para lesiones pigmentadas es la detección de patrones locales, lo que permitirá una evaluación más detallada de dichas lesiones.

En general, esta tesis presenta una amplia gama de técnicas de procesamiento de imágenes para satisfacer diferentes retos biomédicos. Se presenta tanto técnicas simples que en conjunto permiten el desarrollo de un sistema complejo para alcanzar una solución satisfactoria a un problema dado, como nuevas técnicas que abren camino a futuras líneas de investigación.

Publications

International journals.

Sáez A., Rivas E., Montero-Sánchez A., Paradas C., Acha B., Pascual A., Serrano C., Escudero LM. Quantifiable diagnosis of muscular dystrophies and neurogenic atrophies through network analysis. *BMC Medicine.* 11(1), art. no. 77 (2013).

Sáez A., Acha B., Montero-Sánchez A., Rivas E., Escudero LM., Serrano C. Neuromuscular disease classification system. *Journal of Biomedical Optics.* Jun;18(6):066017 (2013).

Sáez A., Mendoza C.S., Acha B., Serrano C. Development and evaluation of perceptually adapted colour gradients. *IET Image Processing* 7 (4) , pp. 355-363 (2013).

Mendoza CS., Perez-Carrasco JA., Sáez A., Acha B., Serrano C. Linearized Multidimensional Earth-Mover's-Distance Gradient Flows. *IEEE transactions on image processing* pp(99), (2013)

Sánchez-Gutiérrez D., Sáez A., Pascual A., Escudero L.M. Topological Progression in Proliferating Epithelia Is Driven by a Unique Variation in Polygon Distribution. *PLOS ONE.* Accepted.

Sáez A., Serrano C., Acha B. Normalized Cut optimization based on color perception findings. A comparative study. *Submitted to Machine Vision and Applications.*

Sáez A., Serrano C., Acha B. Model-based Classification Methods of Global patterns in dermoscopic images. *Submitted to Transactions on Medical Imaging.*

Book chapters.

Sáez A., Serrano C., Acha B. A Review on CAD Tools for Burn Diagnosis. *In Color Medical Image Analysis.* Volume 6, pp 181-202, 2013.

Sáez A., Acha B., Serrano C. Pattern analysis in dermoscopic images. *In Computer Vision Techniques for the Diagnosis of Skin Cancer, published by Springer (Series in BioEngineering)*, 2013.

International conferences.

Fondón I., Van Grinsven M., Sanchez C.I., Sáez A. Perceptually Adapted Method for Optic Disc Detection on Retinal Fundus Images. *26th IEEE International Symposium on Computer-Based Medical Systems* June (2013)

Sánchez-Gutiérrez D., Sáez A., Rivas E., Montero-Sánchez A., Paradas C., Acha B., Pascual A., Serrano C., Escudero L.M. Systems Biology methods for image analysis in development and disease. *Meeting Network Biology SIG of Meeting of the International Society for Computational Biology*. (Best poster award). (2013)

Sáez A., Serrano C., Acha B., Escudero LM. Segmentation of muscle fibres in fluorescence microscopy images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7325 LNCS (PART 2). Aveiro, Portugal. pp. 465-472 (2012).

Sáez A., Fondoń I., Acha B., Jimeńez S., Alemany P., Abbas Q., Serrano C. Optic disc segmentation based on level-set and colour gradients. *6th European Conference on Colour in Graphics, Imaging, and Vision 2012*, CGIV 2012. Amsterdam, the Netherlands. pp. 121-125 (2012).

Sáez A., Acha B., Serrano C. CIELAB based system for burn depth estimation. *6th European Conference on Colour in Graphics, Imaging, and Vision 2012*, CGIV 2012. Amsterdam, the Netherlands. (1), pp. 86-91, (2012).

Sáez A., Serrano C., Acha B. Evaluation perceptual color edge detection algorithms. *5th European Conference on Colour in Graphics, Imaging, and Vision and 12th International Symposium on Multispectral Colour Science 2010*, CGIV 2010/MCS'10 , pp. 222-227 (2010).

Sáez A., Acha B., Serrano C. Segmentation and Classification of Dermatological Lesions. *SPIE International Symposium on Medical Imaging*. San Diego, California (EE.UU.). 76243l-8, (2010).

Sáez A., Acha B., Serrano C. Colorimetric Calibration of Images Captured Under Unknown Illuminants. *11th Congress of the International Colour Association.* AIC 2009. Sidney, Australia. (2009).

Sáez A., Serrano C., Acha B., Gómez Cía PT., Sicilia D. Color Analysis of the Reconstructed Complex Nipple-Areola After a Mastectomy. *11th Congress of the International Colour Association.* AIC 2009. Sidney, Australia. (2009).

National conferences.

Sáez A., Fondón I., Serrano C., Jimenez S., Alemany P., Acha B. Segmentación del disco óptico mediante level-sets con información de color. *XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica.* San Sebastián. (2012).

Sáez A., Serrano C., Acha B. Segmentación perceptiva de imágenes en color basado en corte normalizado. *XXVI Simposio de la Unión Científica Internacional de Radio-URSI.* (2011).

Sáez A., Rivas E., Montero Sánchez A., Paradas C., Acha B., et. al.: Neuronuscular diseases diagnosis through computerized image analysis. *En Centro de Investigación Biomédica en Red, Enfermedades Neurodegenerativas CIBERNED* Diciembre 2011. Donostia (San Sebastian) (2011).

Sáez A., Acha B., Serrano C., Sicilia D. Segmentación y Análisis de Color y Forma del Complejo Areola-Pezón Reconstruido Tras Mastectomía. *XXIII Siposium Nacional de la Unión Científica Internacional de Radio* URSI (2008).

Patents.

Método para obtener información útil para el diagnóstico de enfermedades neuromusculares. Escudero L.M., Montero Sánchez A., Paradas C., Rivas E., Pascual A., Sáez A., Serrano C., Acha B. PCT/ES2012/070796

Bibliography

[Abbas et al., 2013] Abbas, Q., Celebi, M., Serrano, C., Fondón García, I., and Ma, G. (2013). Pattern classification of dermoscopy images: A perceptually uniform model. *Pattern Recognition*, 46(1):86–97.

[Abou-Moustafa et al., 2010] Abou-Moustafa, K., De La Torre, F., and Ferrie, F. (2010). Designing a metric for the difference between gaussian densities. *Advances in Intelligent and Soft Computing*, 83:57–70.

[Ali et al., 2011] Ali, K., Jalil, A., Gull, M., and Fiaz, M. (2011). Medical image segmentation using h-minima transform and region merging technique. In *Proceedings - 2011 9th International Conference on Frontiers of Information Technology, FIT 2011*, pages 127–132.

[Anantha et al., 2004] Anantha, M., Moss, R., and Stoecker, W. (2004). Detection of pigment network in dermatoscopy images using texture analysis. *Computerized Medical Imaging and Graphics*, 28(5):225–234.

[Argenziano et al., 1998] Argenziano, G., Fabbrocini, G., Carli, P., De Giorgi, V., Sammarco, E., and Delfino, M. (1998). Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: Comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Archives of Dermatology*, 134(12):1563–1570.

[Argenziano et al., 2003] Argenziano, G., Soyer, H., Chimenti, S., Talamini, R., Corona, R., Sera, F., Binder, ..., and Kopf, A. (2003). Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the internet. *Journal of the American Academy of Dermatology*, 48(5):679–693.

[Argenziano et al., 2000] Argenziano, G., Soyer, H., and et al. (2000). *Interactive atlas of dermoscopy.* EDRA-Medical Publishing and New Media, Milan.

[Arroyo et al., 2011] Arroyo, J., Zapirain, B., and Zorrilla, A. (2011). Blue-white veil and dark-red patch of pigment pattern recognition in dermoscopic images using

machine-learning techniques. In *IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2011*, pages 196–201.

[Aspland and Shanbhag, 2004] Aspland, J. and Shanbhag, P. (2004). Comparison of color-difference equations for textiles: Cmc(2:1) and ciede2000. *AATCC Review*, 4(6):26–30.

[Barata et al., 2012] Barata, C., Marques, J., and Rozeira, J. (2012). A system for the detection of pigment network in dermoscopy images using directional filters. *IEEE Transactions on Biomedical Engineering*, 59(10):2744–2754.

[Barata et al., 2013] Barata, C., Marques, J. S., and Mendonça, T. (2013). Bag-of-features classification model for the diagnose of melanoma in dermoscopy images using color and texture descriptors. volume ICIAR 2013, pages 547–555.

[Behan et al., 2002] Behan, W., Cossar, D., Madden, H., and McKay, I. (2002). Validation of a simple, rapid, and economical technique for distinguishing type 1 and 2 fibres in fixed and frozen skeletal muscle. *Journal of Clinical Pathology*, 55(5):375–380.

[Bergmeir et al., 2010] Bergmeir, C., García Silvente, M., Esquivias López-Cuervo, J., and Benítez, J. (2010). Segmentation of cervical cell images using mean-shift filtering and morphological operators. In *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, volume 7623.

[Betta et al., 2005] Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., and Scalvenzi, M. (2005). Automated application of the "7-point checklist" diagnosis method for skin lesions: Estimation of chromatic and shape parameters. In *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, volume 3, pages 1818–1822.

[Betta et al., 2006] Betta, G., Di Leo, G., Fabbrocini, G., Paolillo, A., and Sommella, P. (2006). Dermoscopic image-analysis system: Estimation of atypical pigment network and atypical vascular pattern. In *IEEE International Workshop on Medical Measurement and Applications, MeMeA 2006*, volume 2006, pages 63–67.

[Bhattacharyya., 1943] Bhattacharyya., A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 35:99–109.

[Boaventure I., 2009] Boaventure I., G. A. (2009). Method to evaluate the performance of edge detector. In *Proc. of the The Brazilian Symposium on Computer Graphics and Image Processing, Sibgrapi, Brazil*.

[Braun et al., 2005] Braun, R., Rabinovitz, H., Oliviero, M., Kopf, A., and Saurat, J.-H. (2005). Dermoscopy of pigmented skin lesions. *Journal of the American Academy of Dermatology*, 52(1):109–121.

[Capdehourat et al., 2011] Capdehourat, G., Corez, A., Bazzano, A., Alonso, R., and Musé, P. (2011). Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions. *Pattern Recognition Letters*, 32(16):2187–2196.

[Carpenter et al., 1992] Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., and Rosen, D. B. (1992). Fuzzy artmap: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, 3(5):698–713.

[Caselles et al., 1994] Caselles, V., Catté, F., Coll, T., and Dibos, F. (1994). A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31.

[Castillejos et al., 2012] Castillejos, H., Ponomaryov, V., Nino-De-Rivera, L., and Golikov, V. (2012). Wavelet transform fuzzy algorithms for dermoscopic image segmentation. *Computational and Mathematical Methods in Medicine*, 2012.

[Castleman et al., 1984] Castleman, K., Chui, L., Martin, T., and Edgerton, V. (1984). Quantitative muscle biopsy analysis. *Monographs in clinical cytology*, 9:101–116.

[Celebi et al., 2008a] Celebi, M., Iyatomi, H., Stoecker, W., Moss, R., Rabinovitz, H., Argenziano, G., and Soyer, H. (2008a). Automatic detection of blue-white veil and related structures in dermoscopy images. *Computerized Medical Imaging and Graphics*, 32(8):670–677.

[Celebi et al., 2006] Celebi, M., Kingravi, H., Aslandogan, Y., and Stoecker, W. (2006). Detection of blue-white veil areas in dermoscopy images using machine learning techniques. In *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, volume 6144 III.

[Celebi et al., 2008b] Celebi, M., Kingravi, H., Iyatomi, H., Aslandogan, Y., Stoecker, W., Moss, R., Malters, J., Grichnik, J., Marghoob, A., Rabinovitz, H., and Menzies, S. (2008b). Border detection in dermoscopy images using statistical region merging. *Skin Research and Technology*, 14(3):347–353.

[Chan et al., 1996] Chan, S., Leung, K., and Wong, W. (1996). An expert system for the detection of cervical cancer cells using knowledge-based image analyzer. *Artificial Intelligence in Medicine*, 8(1):67–90.

[Chan and Vese, 2001] Chan, T. and Vese, L. (2001). Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.

[Chen et al., 2012] Chen, S., Zhao, M., Wu, G., Yao, C., and Zhang, J. (2012). Recent advances in morphological cell image analysis. *Computational and Mathematical Methods in Medicine*, 2012.

[Cheng et al., 2001] Cheng, H., Jiang, X., Sun, Y., and Wang, J. (2001). Color image segmentation: Advances and prospects. *Pattern Recognition*, 34(12):2259–2281.

[CIE Technical, 2002] CIE Technical, R. (2002). Cie 142-2001, improvement to industrial colour-difference evaluation. *Color Research and Application*, 27(1):61.

[Cross and Jain, 1983] Cross, G. R. and Jain, A. K. (1983). Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-5(1):25–39.

[Cruz-Roa et al., 2011] Cruz-Roa, A., Caicedo, J., and González, F. (2011). Visual pattern mining in histology image collections using bag of features. *Artificial Intelligence in Medicine*, 52(2):91–106.

[Csurka et al., 2004] Csurka, G., Dance, C. R., Fan, L., Willamowski, J., and Bray, C. (2004). Visual categorization with bags of keypoints. In *In Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22.

[Cui and Luo, 2009] Cui, G. and Luo, M. (2009). Testing colour-difference formulae and uniform colour spaces using small colour difference datasets. In *11th Congress of the International Colour Association (AIC), Sidney, Australia*.

[Dalal et al., 2011] Dalal, A., Moss, R., Stanley, R., Stoecker, W., Gupta, K., Calcara, D., Xu, J., Shrestha, B., Drugge, R., Malters, J., and Perry, L. (2011). Concentric decile segmentation of white and hypopigmented areas in dermoscopy images of skin lesions allows discrimination of malignant melanoma. *Computerized Medical Imaging and Graphics*, 35(2):148–154.

[Di Leo et al., 2009] Di Leo, G., Fabbrocini, G., Paolillo, A., Rescigno, O., and Sommella, P. (2009). Towards an automatic diagnosis system for skin lesions: Estimation of blue-whitish veil and regression structures. In *2009 6th International Multi-Conference on Systems, Signals and Devices, SSD 2009*.

[Di Leo et al., 2008] Di Leo, G., Liguori, C., Paolillo, A., and Sommella, P. (2008). An improved procedure for the automatic detection of dermoscopic structures in digital elm images of skin lesions. In *VECIMS 2008 - IEEE Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems Proceedings*, pages 190–195.

[Di Leo et al., 2010a] Di Leo, G., Paolillo, A., Sommella, P., and Fabbrocini, G. (2010a). Automatic diagnosis of melanoma: A software system based on the 7-point check-list. In *Proceedings of the Annual Hawaii International Conference on System Sciences*.

[Di Leo et al., 2010b] Di Leo, G., Paolillo, A., Sommella, P., Fabbrocini, G., and Rescigno, O. (2010b). A software tool for the diagnosis of melanomas automatic implementation of the 7-point check list method. pages 886–891.

[Dowson and Landau, 1982] Dowson, D. and Landau, B. (1982). The fréchet distance between multivariate normal distributions. *Journal of Multivariate Analysis*, 12(3):450–455.

[Doyle et al., 2007] Doyle, S., Hwang, M., Shah, K., Madabhushi, A., Feldman, M., and Tomaszeweski, J. (2007). Automated grading of prostate cancer using architectural and textural image features. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*, pages 1284–1287.

[Dubes and Jain, 1989] Dubes, R. C. and Jain, A. K. (1989). Random field models in image analysis. *Journal of Applied Statistics*, 16(2):131–164.

[Dubowitz et al., 2007] Dubowitz, V., Sewry, C. A., and Lane, R. J. (2007). *Muscle biopsy: a practical approach*. Elsevier Health Sciences.

[Emre Celebi et al., 2007] Emre Celebi, M., Alp Aslandogan, Y., Stoecker, W., Iyatomi, H., Oka, H., and Chen, X. (2007). Unsupervised border detection in dermoscopy images. *Skin Research and Technology*, 13(4):454–462.

[Emre Celebi et al., 2011] Emre Celebi, M., Stoecker, W., and Moss, R. (2011). Advances in skin cancer image analysis. *Computerized Medical Imaging and Graphics*, 35(2):83–84.

[Emre Celebi et al., 2013] Emre Celebi, M., Wen, Q., Hwang, S., Iyatomi, H., and Schaefer, G. (2013). Lesion border detection in dermoscopy images using ensembles of thresholding methods. *Skin Research and Technology*, 19(1):e252–e258.

[Evans and Liu, 2006] Evans, A. and Liu, X. (2006). A morphological gradient approach to color edge detection. *IEEE Transactions on Image Processing*, 15(6):1454–1463.

[Fabbrocinia et al., 2010] Fabbrocinia, G., Betta, G., Leo, G., Liguori, C., Paolillo, A., Pietrosanto, A., Sommella, P., Rescigno, O., Cacciapuoti, S., Pastore, F., Vita, V., and Ayala, I. M. F. (2010). Epiluminescence image processing for melanocytic skin lesion diagnosis based on 7-point check-list: A preliminary discussion on three parameters. *Open Dermatology Journal*, 4(110).

[Fairchild, 2005] Fairchild, M. (2005). *Color Appearance Models*. The Wiley-IS&T Series in Imaging Science and Technology. Wiley.

[Ficarra et al., 2011] Ficarra, E., Di Cataldo, S., Acquaviva, A., and Macii, E. (2011). Automated segmentation of cells with ihc membrane staining. *IEEE Transactions on Biomedical Engineering*, 58(5):1421–1429.

[Fleiss, 1981] Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley, cop1981, New York.

[Fleming et al., 1998] Fleming, M., Steger, C., Zhang, J., Gao, J., Cognetta, A., Pollak, L., and Dyer, C. (1998). Techniques for a structural analysis of dermatoscopic imagery. *Computerized Medical Imaging and Graphics*, 22(5):375–389.

[Frangi et al., 1998] Frangi, R. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention*, pages 130–137.

[Ganster et al., 2001] Ganster, H., Pinz, A., Röhrer, R., Wildling, E., Binder, M., and Kittler, H. (2001). Automated melanoma recognition. *IEEE Transactions on Medical Imaging*, 20(3):233–239.

[Garcia Ugarriza et al., 2009] Garcia Ugarriza, L., Saber, E., Vantaram, S., Amuso, V., Shaw, M., and Bhaskar, R. (2009). Automatic image segmentation by dynamic region growth and multiresolution merging. *IEEE Transactions on Image Processing*, 18(10):2275–2288.

[Garton et al., 2010] Garton, F., Seto, J., North, K., and Yang, N. (2010). Validation of an automated computational method for skeletal muscle fibre morphometry analysis. *Neuromuscular Disorders*, 20(8):540–547.

[Gola Isasi et al., 2011] Gola Isasi, A., García Zapirain, B., and Méndez Zorrilla, A. (2011). Melanomas non-invasive diagnosis application based on the abcd rule and pattern recognition image processing algorithms. *Computers in Biology and Medicine*, 41(9):742–755.

[Grana et al., 2006] Grana, C., Cucchiara, R., Pellacani, G., and Seidenari, S. (2006). Line detection and texture characterization of network patterns. volume 2, pages 275–278.

[Granger, 2008] Granger, E. (2008). A comparison of color difference data and formulas. In *Proceedings of the Technical Association of the Graphic Arts, TAGA*, pages 191–201.

[Greenspan et al., 2004] Greenspan, H., Dvir, G., and Rubner, Y. (2004). Context-dependent segmentation and matching in image databases. *Computer Vision and Image Understanding*, 93(1):86–109.

[Gurcan et al., 2009] Gurcan, M., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., and Yener, B. (2009). Histopathological image analysis: a review. *IEEE reviews in biomedical engineering*, 2:147–171.

[Habekost, 2007] Habekost, M. (2007). Color difference equations and the human eye. In *Proceedings of the Technical Association of the Graphic Arts, TAGA*, pages 291–310.

[Hanmandlu et al., 2008] Hanmandlu, M., Susan, S., Madasu, V., and Lovell, B. (2008). Fuzzy co-clustering of medical images using bacterial foraging.

[Harandi et al., 2010] Harandi, N., Sadri, S., Moghaddam, N., and Amirfattahi, R. (2010). An automated method for segmentation of epithelial cervical cells in images of thinprep. *Journal of Medical Systems*, 34(6):1043–1058.

[Helliwell, 1999] Helliwell, T. (1999). Muscle: Part 1 - normal structure and function. *Current Orthopaedics*, 13(1):33–41.

[Huang and Lai, 2010] Huang, P.-W. and Lai, Y.-H. (2010). Effective segmentation and classification for hcc biopsy images. *Pattern Recognition*, 43(4):1550–1563.

[Ishihara, 1998] Ishihara, S. (1998). *Tests for Colour Blindness*. Hodder Education.

[Iyatomi et al., 2008] Iyatomi, H., Oka, H., Celebi, M., Ogawa, K., Argenziano, G., Soyer, H., Koga, H., Saida, T., Ohara, K., and Tanaka, M. (2008). Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. *Journal of Investigative Dermatology*, 128(8):2049–2054.

[Iyatomi et al., 2007] Iyatomi, H., Oka, H., Celebi, M., Tanaka, M., and Ogawa, K. (2007). Parameterization of dermoscopic findings for the internet-based melanoma screening system. In *Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing, CIISP 2007*, pages 189–193.

[Jain and Zongker, 1997] Jain, A. and Zongker, D. (1997). Feature selection: evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):153–158.

[Julesz, 1981] Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91–97.

[Jung and Kim, 2010] Jung, C. and Kim, C. (2010). Segmenting clustered nuclei using h-minima transform-based marker extraction and contour parameterization. *IEEE Transactions on Biomedical Engineering*, 57(10 PART 2):2600–2604.

[Karen et al., 2009] Karen, P., Ševanec, M., Smerdu, V., Cvetko, E., Kubínová, L., and Eržen, I. (2009). Software for muscle fibre type classification and analysis. *European Journal of Histochemistry*, 53(2):87–95.

[Kashyap and Chellappa, 1983] Kashyap, R. L. and Chellappa, R. (1983). Estimation and choice of neighbors in spatial-interaction models of images. *IEEE Transactions on Information Theory*, IT-29(1):60–72.

[Kass and Witkin, 1987] Kass, M. and Witkin, A. (1987). Analyzing oriented patterns. *Computer Vision, Graphics and Image Processing*, 37(3):362–385.

[Kass et al., 1988] Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331.

[Keinosuke, 1990] Keinosuke, F. (1990). *Introduction to statistical pattern recognition*. Academic Press Professional, Inc., San Diego, CA, USA, 2 edition.

[Khan et al., 2009] Khan, A., Gupta, K., Stanley, R., Stoecker, W., Moss, R., Argenziano, G., Soyer, H., Rabinovitz, H., and Cognetta, A. (2009). Fuzzy logic techniques for blotch feature evaluation in dermoscopy images. *Computerized Medical Imaging and Graphics*, 33(1):50–57.

[Kim et al., 2007] Kim, Y.-J., Brox, T., Feiden, W., and Weickert, J. (2007). Fully automated segmentation and morphometrical analysis of muscle fiber images. *Cytometry Part A*, 71(1):8–15.

[Klemenčič et al., 1998] Klemenčič, A., Kovačič, S., and Pernuš, F. (1998). Automated segmentation of muscle fiber images using active contour models. *Cytometry*, 32(4):317–326.

[Kohavi, 1995] Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *14th international joint conference on Artificial intelligence*, pages 1137–1143.

[Korotkov and Garcia, 2012] Korotkov, K. and Garcia, R. (2012). Computerized analysis of pigmented skin lesions: A review. *Artificial Intelligence in Medicine*, 56(2):69–90.

[Krylov and Nasonov, 2011] Krylov, A. and Nasonov, A. (2011). Edge-directed image interpolation using color gradient information. volume 6979 LNCS of *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*, pages 40–49.

[Kullback, 1997] Kullback, S. (1997). *Information Theory and Statistics*. Dover, New York.

[Lewis, 1998] Lewis, D. D. (1998). Naive (bayes) at forty: The independence assumption in information retrieval. In *Proceedings of the 10th European Conference on Machine Learning*, pages 4–15.

[Li et al., 2005] Li, C., Xu, C., Gui, C., and Fox, M. (2005). Level set evolution without re-initialization: A new variational formulation. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 430–436.

[Li et al., 2010] Li, C., Xu, C., Gui, C., and Fox, M. (2010). Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, 19(12):3243–3254.

[Lucchese and Mitra, 2001] Lucchese, L. and Mitra, S. (2001). Colour segmentation based on separate anisotropic diffusion of chromatic and achromatic channels. *IEE Proceedings: Vision, Image and Signal Processing*, 148(3):141–150.

[Luo et al., 2001] Luo, M., Cui, G., and Rigg, B. (2001). The development of the cie 2000 colour-difference formula: Ciede2000. *Color Research and Application*, 26(5):340–350.

[Madasu and Lovell, 2009] Madasu, V. and Lovell, B. (2009). Blotch detection in pigmented skin lesions using fuzzy co-clustering and texture segmentation. pages 25–31.

[Maglogiannis and Doukas, 2009] Maglogiannis, I. and Doukas, C. (2009). Overview of advanced computer vision systems for skin lesions characterization. *IEEE Transactions on Information Technology in Biomedicine*, 13(5):721–733.

[Malladi et al., 1995a] Malladi, R., Sethian, J. A., and Vemuri, B. C. (1995a). Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175.

[Malladi et al., 1995b] Malladi, R., Sethian, J. A., and Vemuri, B. C. (1995b). Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175.

[Mandic et al., 2006] Mandic, L., Grgic, S., and Grgic, M. (2006). Comparison of color difference equations. In *Proceedings Elmar - International Symposium Electronics in Marine*, pages 107–110.

[Manjunath and Chellappa, 1991] Manjunath, B. and Chellappa, R. (1991). Unsupervised texture segmentation using markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):478–482.

[Mclachlan and Peel, 2000] Mclachlan, G. and Peel, D. (2000). *Finite Mixture Models*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1 edition.

[McLaren, 1976] McLaren, K. (1976). Development of the cie 1976 (l a b) uniform colour space and colour-difference formula.

[Mendoza et al., 2009] Mendoza, C., Serrano, C., and Acha, B. (2009). Pattern analysis of dermoscopic images based on fscm color markov random fields. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5807 LNCS:676–685.

[Meunier et al., 2010] Meunier, B., Picard, B., Astruc, T., and Labas, R. (2010). Development of image analysis tool for the classification of muscle fibre type using immunohistochemical staining. *Histochemistry and Cell Biology*, 134(3):307–317.

[Meyer and Beucher, 1990] Meyer, F. and Beucher, S. (1990). Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1):21–46.

[Mirzaalian et al., 2012] Mirzaalian, H., Lee, T., and Hamarneh, G. (2012). Learning features for streak detection in dermoscopic color images using localized radial flux of principal intensity curvature. pages 97–101.

[Moreno et al., 2010] Moreno, R., Graña, M., and D'Anjou, A. (2010). An image color gradient preserving color constancy.

[Nourmohamadi and Pourghassem, 2012] Nourmohamadi, M. and Pourghassem, H. (2012). Dermoscopy image segmentation using a modified level set algorithm. In *Proceedings - 4th International Conference on Computational Intelligence and Communication Networks, CICN 2012*, pages 286–290.

[Ojala et al., 2002] Ojala, T., Pietikäinen, M., and Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987.

[Osher and Fedkiw, 2003] Osher, S. and Fedkiw, R. (2003). *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences. Springer.

[Osher and Sethian, 1988] Osher, S. and Sethian, J. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12–49.

[Pehamberger et al., 1987] Pehamberger, H., Steiner, A., and Wolff, K. (1987). In vivo epiluminescence microscopy of pigmented skin lesions. i. pattern analysis of pigmented skin lesions. *Journal of the American Academy of Dermatology*, 17(4):571–583.

[Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639.

[Plataniotis and Venetsanopoulos, 2000] Plataniotis, K. N. and Venetsanopoulos, A. N. (2000). *Color image processing and applications*. Springer-Verlag New York, Inc., New York, NY, USA.

[Plissiti et al., 2011] Plissiti, M., Nikou, C., and Charchanti, A. (2011). Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. *IEEE Transactions on Information Technology in Biomedicine*, 15(2):233–241.

[Pratt, 1978] Pratt, W. (1978). *Digital image processing*. Wiley-interscience publication. Wiley.

[Pudil et al., 1994] Pudil, P., Novovičová, J., and Kittler, J. (1994). Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125.

[Rangayyan et al., 2011] Rangayyan, R., Acha, B., and Serrano, C. (2011). *Color Image Processing With Biomedical Applications*. Press monograph 206. SPIE Press.

[Refaeilzadeh et al., 2009] Refaeilzadeh, P., Tang, L., and Liu, H. (2009). Cross-validation. In Liu, L. and Özsu, M. T., editors, *Encyclopedia of Database Systems*, pages 532–538. Springer US.

[Report, 1995] Report, C. T. (1995). *Industrial color difference evaluation*. CIE Publication 116:1995, Central Bureau, Vienna.

[Rezze et al., 2006] Rezze, G., De Sá, B., and Neves, R. (2006). Dermoscopy: The pattern analysis. *Anais Brasileiros de Dermatologia*, 81(3):261–268.

[Robertson, 1978] Robertson, A. R. (1978). Cie guidelines for coordinated research on colour-difference evaluation. *Color Research and Application*, 3(3):149–151.

[Romaní et al., 2010] Romaní, S., Prados-Suárez, B., Sobrevilla, P., and Montseny, E. (2010). Cytoplasm contour approximation based on color fuzzy sets and color gradient. volume 6178 LNAI of *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*, pages 645–654.

[Rubner et al., 2000] Rubner, Y., Tomasi, C., and Guibas, L. (2000). Earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121.

[Sadeghi et al., 2012a] Sadeghi, M., Lee, T., McLean, D., Lui, H., and Atkins, M. (2012a). Global pattern analysis and classification of dermoscopic images using textons. In *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, volume 8314.

[Sadeghi et al., 2012b] Sadeghi, M., Lee, T., McLean, D., Lui, H., and Atkins, M. (2012b). Oriented pattern analysis for streak detection in dermoscopy images. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 15(Pt 1):298–306.

[Sadeghi et al., 2013] Sadeghi, M., Lee, T., McLean, D., Lui, H., and Atkins, M. (2013). Detection and analysis of irregular streaks in dermoscopic images of skin lesions. *IEEE Transactions on Medical Imaging*, 32(5):849–861.

[Sadeghi et al., 2011] Sadeghi, M., Razmara, M., Lee, T., and Atkins, M. (2011). A novel method for detection of pigment network in dermoscopic images using graphs. *Computerized Medical Imaging and Graphics*, 35(2):137–143.

[Sadeghi et al., 2010] Sadeghi, M., Razmara, M., Wighton, P., Lee, T., and Atkins, M. (2010). Modeling the dermoscopic structure pigment network using a clinically inspired feature set. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 6326 LNCS, pages 467–474.

[Sadri et al., 2013] Sadri, A., Zekri, M., Sadri, S., Gheissari, N., Mokhtari, M., and Kolahdouzan, F. (2013). Segmentation of dermoscopy images using wavelet networks. *IEEE Transactions on Biomedical Engineering*, 60(4):1134–1141.

[Schmid, 1999] Schmid, P. (1999). Segmentation of digitized dermatoscopic images by two-dimensional color clustering. *IEEE Transactions on Medical Imaging*, 18(2):164–171.

[Serrano and Acha, 2009] Serrano, C. and Acha, B. (2009). Pattern analysis of dermoscopic images based on markov random fields. *Pattern Recognition*, 42(6):1052–1057.

[Sertel et al., 2011] Sertel, O., Dogdas, B., Chiu, C., and Gurcan, M. (2011). Microscopic image analysis for quantitative characterization of muscle fiber type composition. *Computerized Medical Imaging and Graphics*, 35(7-8):616–628.

[Sethian, 1999] Sethian, J. (1999). *Level Set Methods and Fast Marching Methods*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press.

[Sfikas et al., 2005] Sfikas, G., Constantinopoulos, C., Likas, A., and Galatsanos, N. (2005). An analytic distance metric for gaussian mixture models with application in image retrieval. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3697 LNCS:835–840.

[Shen and Berns, 2011] Shen, S. and Berns, R. (2011). Color-difference formula performance for several datasets of small color differences based on visual uncertainty. *Color Research and Application*, 36(1):15–26.

[Shrestha et al., 2010] Shrestha, B., Bishop, J., Kam, K., Chen, X., Moss, R., Stoecker, W., Umbaugh, S., Stanley, R., Celebi, M., Marghoob, A., Argenziano, G., and Soyer, H. (2010). Detection of atypical texture features in early malignant melanoma. *Skin Research and Technology*, 16(1):60–65.

[Situ et al., 2010] Situ, N., Wadhawan, T., Yuan, X., and Zouridakis, G. (2010). Modeling spatial relation in skin lesion images by the graph walk kernel. volume 2010, pages 6130–6133.

[Situ et al., 2008] Situ, N., Yuan, X., Chen, J., Zouridakis, G., and Yuan, X. (2008). Malignant melanoma detection by bag-of-features classification. pages 3110–3113.

[Skrøvseth et al., 2010] Skrøvseth, S., Schopf, T., Thon, K., Zortea, M., Geilhufe, M., Møllersen, K., Kirchesch, H., and Godtliebsen, F. (2010). A computer aided diagnostic system for malignant melanomas. In *2010 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies, ISABEL 2010*.

[Soille, 2003] Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2 edition.

[Steger, 1998] Steger, G. (1998). An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(2):113–125.

[Stoecker et al., 2005] Stoecker, W., Gupta, K., Stanley, R., Moss, R., and Shrestha, B. (2005). Detection of asymmetric blotches (asymmetric structureless areas) in dermoscopy images of malignant melanoma using relative color. *Skin Research and Technology*, 11(3):179–184.

[Stoecker and Moss, 1992] Stoecker, W. and Moss, R. (1992). Editorial: digital imaging in dermatology. *Computerized Medical Imaging and Graphics*, 16(3):145–150.

[Tanaka et al., 2008] Tanaka, T., Torii, S., Kabuta, I., Shimizu, K., and Tanaka, M. (2008). Pattern classification of nevus with texture analysis. *IEEJ Transactions on Electrical and Electronic Engineering*, 3(1):143–150.

[Tanaka et al., 2004] Tanaka, T., Torii, S., Kabuta, I., Shimizu, K., Tanaka, M., and Oka, H. (2004). Pattern classification of nevus with texture analysis. In *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, volume 26 II, pages 1459–1462.

[Todman and Claridge, 2000] Todman, A. and Claridge, E. (2000). Low-level grouping mechanisms for contour completion. *Information Sciences*, 125(1-4):19–35.

[Tommasi et al., 2008] Tommasi, T., Orabona, F., and Caputo, B. (2008). Discriminative cue integration for medical image annotation. *Pattern Recognition Letters*, 29(15):1996–2002.

[Tuceryan and Jain., 1998] Tuceryan, M. and Jain., A. (1998). Texture analysis. In Chen, C. H., Pau, L. F., and Wang, P. S. P., editors, *Handbook of Pattern Recognition and Vision*. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2nd edition.

[Varma and Zisserman, 2003] Varma, M. and Zisserman, A. (2003). Texture classification: Are filter banks necessary? volume 2, pages II/691–II/698.

[Wang et al., 2012] Wang, H., Cui, G., Luo, M., and Xu, H. (2012). Evaluation of colour-difference formulae for different colour-difference magnitudes. *Color Research and Application*, 37(5):316–325.

[Wang et al., 2013] Wang, J.-Y., Bensmail, H., and Gao, X. (2013). Joint learning and weighting of visual vocabulary for bag-of-feature based tissue classification. *Pattern Recognition*.

[Wesolkowski et al., 2000] Wesolkowski, S., Jernigan, M., and Dony, R. (2000). Comparison of color image edge detectors in multiple color spaces. In *IEEE International Conference on Image Processing*, volume 2, pages 796–799.

[Wighton et al., 2011] Wighton, P., Lee, T., Lui, H., McLean, D., and Atkins, M. (2011). Generalizing common tasks in automated skin lesion diagnosis. *IEEE Transactions on Information Technology in Biomedicine*, 15(4):622–629.

[Wighton et al., 2009] Wighton, P., Sadeghi, M., Lee, T., and Atkins, M. (2009). A fully automatic random walker segmentation for skin lesions in a supervised setting. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5762 LNCS(PART 2):1108–1115.

[Xia et al., 2006] Xia, Y., Feng, D., and Zhao, R. (2006). Adaptive segmentation of textured images by using the coupled markov random field model. *IEEE Transactions on Image Processing*, 15(11):3559–3566.

[Xue-Wei and Zhang, 2008] Xue-Wei, L. and Zhang, X.-R. (2008). A perceptual color edge detection algorithm. In *Proceedings - International Conference on Computer Science and Software Engineering, CSSE 2008*, volume 1, pages 297–300.

[Yüksel and Borlu, 2009] Yüksel, M. and Borlu, M. (2009). Accurate segmentation of dermoscopic images by image thresholding based on type-2 fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 17(4):976–982.

[Yoshino et al., 2004] Yoshino, S., Tanaka, T., Tanaka, M., and Oka, H. (2004). Application of morphology for detection of dots in tumor. In *Proceedings of the SICE Annual Conference*, pages 407–410.

[Zhou et al., 2011] Zhou, H., Schaefer, G., Celebi, M., Lin, F., and Liu, T. (2011). Gradient vector flow with mean shift for skin lesion segmentation. *Computerized Medical Imaging and Graphics*, 35(2):121–127.

[Zhu et al., 1999] Zhu, S.-Y., Plataniotis, K., and Venetsanopoulos, A. (1999). Comprehensive analysis of edge detection in color image processing. *Optical Engineering*, 38(4):612–625.

List of Figures

List of Tables

Glossary

$L^*a^*b^*$ Uniform colour space defined by CIE in 1976. 4

VD Colour vectorial gradient proposed in section 3.2.2. 46

APN Atypical pigment network. 68

BoF Bag of features. 93

C Control (no disease). 25, 29, 32

CAD Computer aided diagnosis. 3, 4, 12, 55, 106

CIE Commission internationale de l'éclairage (International Commission on Illumination). 4, 48

CIE94 Colour difference equation defined by CIE in 1994. 39, 40, 49, 56, 106

CIEDE2000 Colour difference equation defined by CIE in 2000. 40, 49, 56, 106

CIELAB Euclidean distance computed in $L^*a^*b^*$ colour space. 40, 48, 49, 56, 106

EMD Earth movers distance. 89

FSCM Finite symmetric conditional model. 88

GMM Gaussian mixture model. 92

KL Kullback Leibler. 89

KNN K-Nearest Neighbour algorithm. 91

LSF Level set function. 44, 46

MD Muscular dystrophy. 11, 25, 29, 32

MND Neuromuscular diseases. 9

MRF Markov random fields. 87, 89, 93

NA Neurogenic atrophy. 11, 25, 29, 32

PCA Principal component analysis. 30, 33, 106

PDE Partial differential equation. 44

ROI Region of interest. 23

SBS sequential backward selection. 25

SFS Sequential forward selection. 25